UNIVERSIDAD CARLOS III DE MADRID

Department of Signal Theory and Communications

DOCTORAL THESIS

STRESS LEVEL ASSESSMENT WITH NON-INTRUSIVE SENSORS

Author: FRANCISCO HERNANDO GALLEGO

Supervised by: ANTONIO ARTÉS RODRÍGUEZ

TBD, 2018

Tesis Doctoral:	STRESS LEVEL ASSESSMENT WITH NON-INTRUSIVE SENSORS
Autor:	Francisco Hernando Gallego
Director:	D. Antonio Artés Rodríguez
Fecha:	TBD, 2018

Tribunal

Presidente: TBD

Vocal: TBD

Secretario: TBD

Money makes the world go round.

legal or illegal,

good guys or bad guys,

we all chase money.

Acknowledgements

I would like to acknowledge some people who have contributed the most to the successful completion of this work.

First of all, I want to express my gratitude to Antonio Artés for his guidance. You have not only enabled and supported this thesis in all possible ways but have also taught me how to become a researcher, collaborator, and human being. I also want to acknowledge Fernando Pérez, who has giving me an opportunity to enroll an internship in Switzerland focusing on sport betting and learn new ways of thinking. Both will always be a role model to follow.

I would like to thank everyone in our research group: Fran I, Pablo G., Isa, Melanie, Alfred, Paloma, Alberto, Vivian, Deniz, Gonzalo V., Gonzalo R., Tobi, David R., Pablo O., Victor E., Gesús F., Pablo M., Juanjo, Alejandro, Aurora and Sara, for making this time really enjoyable. Thanks for all the laughs and good vibes. Grace, it has been a huge pleasure to share all our experiences in the best 'cubiculo'. Thanks also to my Bachelor Degree students: Sergio, Verónica and Sara, part of this thesis is yours.

Special thanks to my hometown friends, 'Las ranas': Álvaro, Diego, Néstor, Pablo, Sevi, Ribete and Ruso, for dealing with me at my best and worst, always supporting and encouraging me. Without all of you, I would not be writing now.

I would like to thank my friends from 'Valseca', always standing alongside me. Pablo M., Cuqui, Mónica and Maru, do not change ever, please. Pablete, you have also become a Ph.D. photoshop-maker, ¡congrats! my right-hand man.

Joe & Nacho, thanks for day after day analyzing and discussing sport bets. And also thanks for being always supporting me. Thanks will never suffice.

Finally, the achievement of this doctoral thesis would have never been possible without the unwavering support of my family and beloved ones. Thank you to my parents and sisters, Maxi & Paco, Ana & Mat and Neli, I owe you more than I could possibly express with words. 'Compañeros de vida, compañeros de ilusión'.

"If you want to go fast, go alone. If you want to go far, go together."

Abstract

Stress is an involuntary reaction where the human body changes from a calm state to an excited state in order to preserve the integrity of the organism. Small amount of stress should be good to became entrepreneur and learn new ways of thinking, but continuous stress can carry an array of daily risks, such as, cardiovascular diseases, hair loss, diabetes or immune dysregulation. Recognize how, when and where it occurs has become a step in stress assessment.

Stress recognition starts from 1973 until now. This disease has become a problem in recent years because has increased the number of cases, especially in workers where his/her performance decreases. Stress reactions are provoked for the Autonomous Nervous System (ANS) and one way to estimate it could be found in physiological signals. A list of a variety wearable sensor is presented to capture these reactions, trying to minimize the risk of distraction due to external factors.

The aim of this work thesis is to detect stress for level assessment. A combination of different physiological signals is selected to extract stress feature an classify in a rating scale from relax to breakdown situations.

This thesis proposes a new feature extraction model to understand physiological Galvanic Skin Response (GSR) reactions. Last methods conclude in incongruent results that are not interpretable. This model propose a robust algorithm that can be used in real-time (low time computability) and results are sparse in time to obtain an easily statistical and graphical interpretation.

Signal processing methods of heart rhythm and hormone cortisol are included to develop a robust feature extraction method of stress reactions. A combination of electrodermal, heart and hormone analysis is presented to know in real-time the state of the individual. These features have been selected because the acquisition is non-intrusive avoiding other factor such as distractions.

This thesis is application-focused and highly multidisciplinary. A complete feature extraction model is presented including the new electrodermal model named and usual heart rhythm techniques. Three experiments were evaluated: a) a fea-

ture selection model using neurocognitive games, b) a stress classifier in time during public talks, and c) a real-time stress assessment classifier in a five-star rating scale.

This thesis improve stress detection overcoming a system to capture physiological responses, analyze and conclude a stress assessment decision. We discussed past state of the art and propose a new method of feature extraction using signal processing improvements. Three different scenarios were evaluated to confirm the achievement of aims proposed.

Resumen

TBD

Contents

List of Acronyms

ANS	Autonomous Nervous System
ARMA	Autoregressive–Moving-Average Model
AWGN	Additive White Gaussian Noise
BVP	Blood Volume Presure
CDA	Continuous Decomposition Analysis
CNS	Central Nervous System
ECG	Electro Cardiogram
ECG	Electroencephalogram
EDA	Electrodermal Activity
GSR	Galvanic Skin Response
HRV	Heart Rate Variability
LARS	Least Angle Regression
LASSO	Least Absolute Shrinkage and Selection Operator
LDA	Linear Discriminant Analysis
LFPC	Log Frequency Power Coefficients

LOO	Leave-One-Out
MSE	Mean Square Error
NS	Nervous System
PNS	Parasympathetic Nervous System
PPG	Photoplethysmography
PS	Parasympathetic System
RMSSD	Root Mean Square of Successive Differences
RR	Inter-Beat
SCL	Skin Conductance Lelvel
SCR	Skin Conductance Response
SDNN	Standard Deviation of NN Intervals
SNS	Sympathetic Nervous System
SS	Sympathetic System
SVM	Support Vector Machine
TSST	Trier Social Stress Test
UAV	Unmanned Aerial Vehicle
US	United States

1 Introduction

1.1 Motivation

Stress is a body reaction that could appear in our daily life everyday. Do you remember the last time you felt stressed? Maybe you had to giving a talk and very little time to prepare it or perhaps you have a deadline and you do not have time to finish your document and send it properly. Although you might not have been completely aware about feeling stressed, your body might have been experiencing a sequence of physiological changes that may have induced such as pupil dilation, deeper breathing, intensified beating of the heart, or increased muscle tension, among many other possible changes. These physiological alterations and their associated behavioral effects, plays a role in our daily functioning [14]. Stress not only regulates processes such as attention and memory acquisition but also helps us tune the body to face daily challenges and threats [53]. Some of this response

during daily activity can result in long-term stress, contributing to a wide array of risks, including cardiovascular disease, cerebrovascular disease, diabetes, and immune dysregulation [5].

The body experiences a series of physiological events driven by the two branches of the ANS in a stressful situation. The Sympathetic Nervous System (SNS) mobilizes the body's resources in response to a challenge or a threat (e.g., quickens the pulse, deepens the respiration and tenses the muscles) and the Parasympathetic Nervous System (PNS) works antagonistically to control this process. During this process, people may experience enhanced arousal, improved cognitive functioning and concentration, increased ability to withstand pain, and accelerated motor reflexes, preparing the body to face life-threatening situations. This type of instantaneous stress (usually named short-term stress) has been defined as a reaction from a calm state to an excited state in order to preserve the integrity of the organism [25]. It has been estimated that the business costs associated with stress are around $300 billion per year in the United States (US) alone [64]. Despite the well-studied negative outcomes, stress is still considered a necessary evil by many people as it helps us keep up with the pace of modem society.

This thesis focus on short-term stress changes that can be triggered in a few of seconds instead of focusing on long-term stress. Continuous short-time stress can impair decision making, decrease productivity, and lead to high amounts of work accidents [52]. The relationship between arousal and performance plays a role in terms of stress appraisal and health outcomes. According to the Yerkes-Dodson Law [84], this achievement is a function of arousal, in Figure 1.1, performance increases with arousal when the individual feel relaxed, then reaches its peak at the highest arousal level because the subject is involved in the task, and decreases when the individual feel in a breakdown or anxiety situation.

We already know how humans reacts to stress situations, next question is how can we measure this short-term stress in the least non-intrusive way possible. Researchers have studied a wide variety of approaches, such as self-reports to interpret the performance, hormone analysis and the measurement of physiological signals to

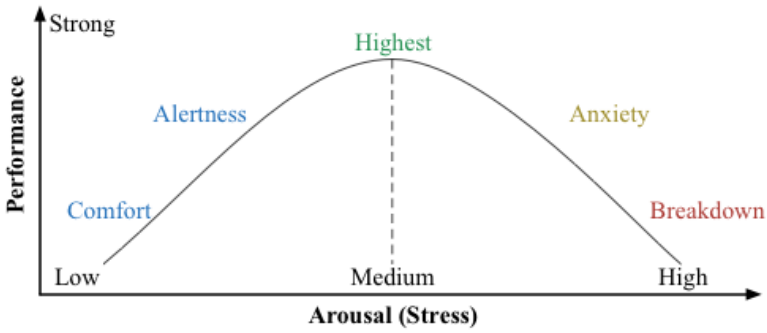


Figure 1.1: Yerkes-Dodson Law: relation between arousal and performance

evaluate arousal. However, each of these approaches has its own set of limitations such as requiring the attention of the person, being obtrusive and/or capturing partial information about stress. Moreover, there is often great variability in how people perceive, experience and physiologically express stress, obstructing efforts to build a stress recognition system.

We propose two points to model the Yerkes & Dodson curve while an individual is performing a specific test:

1. x-axis can be expressed as the arousal (in this case, the short-term stress). The x-axis represents in time, the stress of the subject. The individual is elicit with in a controlled environment from a comfort state to an breakdown situation,

2. and y-axis is the performance obtained for this individual during this test.

Emotional states provoke changes in different physiological signals that can be measured in order to obtain information about the mental state of the individual. Several physiological signals in different status are already analyzed in the literature. Electrodermal activity varies significantly between stress and relax situations and can be extracted parameters of attention an short-term stress from the skin monitoring the GSR. Hearth rhythm varies comparing between situations and two signals can be monitored: Electro Cardiogram (ECG) and Blood Volume Presure (BVP). Also another signals such as speech, hormone cortisol, Electroencephalo-

gram (ECG) or pupil diameter can be biomarker patters associated with disease stage, stress, as well as treatment efficacy.

A system has to be developed once are known possible signals to find stress. Wearable sensors are increasingly taking part in daily activities, not only because of the recent society health concern, but also due to their relevance in the medical industry. A wide variety of sensors can monitor physiological signals, but, Which is are the requirements to use one or other?. First of all, the sensor should be non-intrusive to avoid disturbing and also because the person who wears it could be uncomfortable. Secondly, the system acquisition should be via wireless technology instead of via link assuming a loss of quality. Via link is intrusive and depends of the length of the link. Finally, the sampling rate of the acquisition and the material of the sensor have to be revised. A large majority or the sensors provide signals with artifacts and out-layers that complicate the signal processing addressed. The x-axis scale can be modeled using the reactions extraxted from these signals.

The second problem to overcome is to fit the y-axis of the Yerkes & Dodson curve. For over a century, psychologists and behavioral scientists have thoroughly studied stress and other emotions in laboratory settings in a subjective way. However, these emotions should be partly contrived in a objective way -i.e. elicited by controlled stimuli-, which may not matter to the participants, at least not compared with real life stressful events. With the recent improvement of wearable sensors, researchers have started to study more natural and spontaneous emotions and their role in real-life interactions, creating a new set of complex challenges that need to be addressed. These methods have been traditionally tested in controlled or semi-controlled settings where most of the real-life variables that introduce noise are controlled or eliminated. The research at the core of this thesis combines state-of-the-art physiological signals with machine learning methods to advance the measurement and analysis of stress at work.

The initiative of this thesis was proposed for *AIRBUS Group* company. They create a project entitled *SAVIER*, which is a real demonstrator to make operators of Unmanned Aerial Vehicle (UAV) work easier. The project was divided into

twelve thesis such as speech recognition, multi-UAV coordination, gesture control, etc, and one of them was this work of stress level assessment. This kind of workers are put under strong pressure while are achieving a mission. The effectiveness of these operators depends on their stress level, so it would be useful to be able to monitor and evaluate this magnitude. This project explains what stress is, how it can be measured in non-intrusive way and analyze for determine a stress level in a scale ranking.

1.2 Scientific Aims and Perspectives

The specific research aims of the thesis are the following:

- Create a novel of physiological feature extraction parameters that can be unobtrusively measure with non-intrusive sensors.
- Evaluation of different physiological features to know which parameters are more relevant in stress assessment.
- Interpretation of the validity of these novel parameters in controlled scenarios.
- Assessment of generalization of the previous results in a longitudinal real-life workplace settings.
- Development a software algorithm to automatically classify the level of stress in a five-star rating scale based on physiological reactions.

1.3 Contributions

This thesis is multidisciplinary, bringing an analysis to deal with real-world stress problems in the fields of public talks, individual characterization and neurocognitive games. Throughout this thesis, we address the following points:

(A) A new feature extraction model of electrodermal sparse reactions.

(B) A novel of the most relevant physiological features to detect stress reactions modeled in objective way using wearable sensors.

(C) A new real-time method to characterize and classify stress levels in a five-star rating scale.

The contributions of this thesis have also been or will be partially published in [33, 35, 34, 36]. This thesis corresponds to a complete robust stress feature extraction algorithm applied to different relevant situations where stress may be controlled. We summarize our contributions below.

1.3.1 Electrodermal Feature Extraction Model

A GSR extraction technique has been developed in order to interpret Electrodermal Activity (EDA) records, which can be useful both for ambulatory and health applications. The core of the proposed approach is a feature extraction scheme that is based on a non-negative sparse deconvolution of the observed GSR signals. Unlike previous approaches, the resulting algorithm is fast (immediately extracting the skin conductance level and response), efficient (being able to work with any sampling rate and signal length), and highly interpretable (due to the sparsity of the extracted phasic component of the GSR).

1.3.2 Wearable Sensors Used to Detect Stress Levels

This contribution presents a controlled experiment design to know which physiological features are more relevant in human decisions. During the experiments, subjects are requested to either play a neurocognitive game using a computer, or relax during interleaved intervals of time. Nine subjects performed the experiments twice and twenty four once to train the model. The main objective is to analyze the capability of the extracted features to understand individual behaviors. A binary classification problem is proposed to determine whether a person was playing or relaxing, achieving 13.31% mean error.

1.3.3 Classification Physiological Stress Features

This study proposes a new framework to process signals while an individual is discussing a public talk trying to classify his/her stress levels. A dataset is presented, composed of 17 one hour talks where speech, electrodermal activity and heart rhythm are recorder using non-intrusive sensors. The proposed method acquire 12 features each minute and classify in three states: pre-presentation, during talk and questions time. A database of 9 speeches were used as training set and the rest of 8 talks as test data. We show that the proposed framework makes it possible to detect distinctive stress patterns, verifying an average of 15.05% percentage of error between the three classes, achieving a model to use in stress detection.

1.4 Thesis Outline

The outline of the remainder of the thesis is as follows: Chapter 2 overviews relevant research in the context of stress measurement and some of its main challenges. In particular, the chapter describes: 1) some of the commonly used workplace stress models, 2) several approaches to measuring stress, 3) different approaches to comfortably measure physiological signals, 4) the most important physiological signals used in stress recognition. Chapter 3 provides new method of feature extraction for GSR signals. The model is completely explained and compared with other methods developed in the past. Chapter 4 presents some feature extraction techniques for different human activities. It is divided into heart rhythm, electrodermal activity, speech and hormone analysis. Chapter 5 discuses three experiment developed to elicit stress in a controlled environment to understand the stress reactions of the participants. Lastly, Chapter 6 concludes the dissertation of this thesis and finished with some possible future lines.

2

Stress Assessment

This chapter provides an overview of relevant research in the context of stress measurement and some of its main challenges. In particular, the chapter describes: 1) definition and different types of stress, 2) how the Autonomous Nervous System (ANS) reacts from stress situations, 3) objective versus subjective stress measurements, 4) a novel of the most relevant physiological signals used in stress assessment, 5) a list of possible wearable sensors, 6) and finally some experiments of stress recognition using physiological signals.

2.1 Stress Definition

Stress is a natural physical and mental reaction to life experiences. Everyone expresses stress from time to time. Anything from everyday responsibilities like work and family to serious life events such as a new diagnosis, war, or the death

of a loved one can trigger stress. For immediate, short-term situations, stress can be beneficial to health. The body responds to stress by releasing hormones that increase heart and breathing rates, and ready your muscles to respond. This is called fight-or-flight response that the body triggers in times of duress.

But stress is meant to be temporary. The body should return to a natural state after the situation has passed. Heart rate should slow, muscles should relax, and breathing should return to normal. The pressures and demands of modern life may put the body in a heightened state for a long period of time, making heart pump hard and your blood vessels constrict for longer than body can handle. Over time, these physiological demands can take a toll on body.

Yet if stress response does not stop firing, and these stress levels stay elevated far longer than is necessary for survival, it can take a toll on health. Long-term stress can cause a variety of symptoms and affect the overall well-being. This chain of physiological changes and their associated behavioral effects, plays a role in our daily functioning. This thesis borrows the definitions of demands in these three cases:

Long-term stress

It is body's immediate reaction to a new challenge, event, or demand, and it triggers your fight-or-flight response. As the pressures of a near-miss automobile accident, an argument with a family member, or a costly mistake at work sink in, body turns on this biological response.

Short-term stress

It comes about as the result of a situation that has not been resolved or continued for many years prior to being resolved. This might be a traumatic event that happened during childhood. Although resolved, the feelings surrounding the situation may not have been dealt with and chronic stress remains. There may also be an ongoing situation, such as family abuse, dysfunctional home or an ongoing illness in the family. This stress has the ability to create additional health problems, for example heart disease or stomach ulcers.

Individual performance

Every person has different responses in presence of stress situations, but present similar behaviors performing the same activity. Performance achieved could have variance as describe [83], i.e. the time reaction for each individual carrying out a task. It is important to characterize the capacity and abilities of each person due their variance.

One of the main challenges when defining and measuring stress is the great variability in how people perceive and experience stress. Stress reactions are different for each person, but present similar behaviors doing the same exercise. Moreover, this same activity may elicit different levels of stress in the same person from one day to another, perhaps depending on factors such as the amount of sleep s/he had during the previous day [59], whether s/he exercises regularly [53], and the amount and type of social interactions s/he had during the day [50]. All of these factors complicate the task of estimate stress.

2.2 Autonomous Nervous System (ANS)

There are stressful situations where individuals must deal with changes from a calm state to an excited state. But, how does the body react to these stimulus? Why are these changes in the body caused? The answer is given by the Sympathetic System (SS) and the Parasympathetic System (PS), which are part of the Nervous System (NS) of the human body. NS is divided into two components [62]: Central Nervous System (CNS) and ANS. This thesis will focus on the ANS, because it is the system that controls involuntary actions of the human body, such as the beating of heart, sweating or digestion. A hierarchical scheme of the NS is shown in Figure 2.1.

ANS is an efferent system (i.e. it carries impulses from CNS to peripheral organs) and it can be divided into two subsystems: Sympathetic Nervous System (SNS) and Parasympathetic Nervous System (PNS). Both subsystems make actions that can seem opposite each other. The SS acts in urgent cases causing

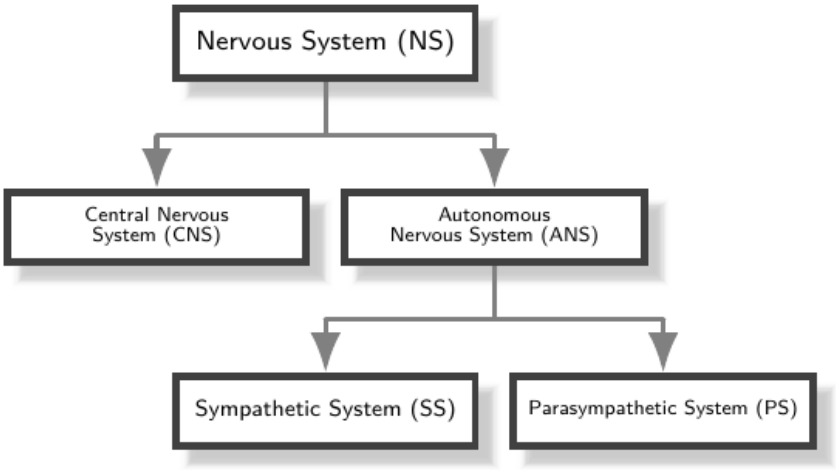


Figure 2.1: Hierarchical scheme of the Nervous System (NS).

reactions such as the fact of accelerating the pulse rhythm or the breathing or breaking the digestion. This subsystem acts on this way to prepare people in order to use their maximum energy. However, the PS keeps energy in order to maintain the properly operation of the human body after these urgent cases where the ANS acts.

Both subsystems control different functions of the human body which are sensitive to emotional states [61], and emotions are related to stress because people who cannot or do not know how to deal with these emotional situations are more likely to suffer from this stress problem. Emotions can be elicited by perception, imagery, anticipation and action [12]. There have been several researches where different stimuli have been applied aiming at eliciting different emotions, and it is just in this point where can take into account a set of different physiological signals that can be measured in a non-intrusive way, and thanks to them, it is possible to obtain information of the emotional state.

2.3 Objective vs Subjective Stress Measurement

Measuring stress has been the focus of interest of psychological and psychophysiological researchers for many decades. Being able to automatically quantify stress

during daily life could help people not only to better understand what events elicited the highest stress levels during their daily activity but also to prevent the negative outcomes associated with long-term stress.

From the point of view of psychology, the standard of stress measurement is self-report measures that can be collected through retrospective surveys and/or experience sampling. The psychologists consider that the main advantage of this approach is that it is subjective and, therefore, it can potentially capture the individual personal experience of stress. There are a wide array of surveys to quantify different types of stress and their frequency during daily life. For instance, the Daily Stress Inventory [13] allows one to quantify the daily stress by counting the number and type of stress that appeared throughout the day. While retrospective surveys enable capturing very detailed information, they are negatively affected by recall problems. On the other hand, experience sampling methods minimize these problems but severely limit the amount of information that can be captured. A common limitation is that they require the full attention of the person, which can be disruptive when considering very frequent real-life measurements. Moreover, they assume that the monitored person can appropriately identify and express their own emotions, and that they are willing to do so. These assumptions are not always accurate in real-life (e.g., people with emotional impairments, excessive work overload that negatively impacts accurate self-reflection, or discomfort reporting some experiences).

The main alternative to self-report measures are objective measures. For instance, information about stress responses can also be obtained by analyzing hormones such as cortisol or adrenaline that can be gathered from saliva and blood samples. However, these measures commonly entail costly and slow analysis. An alternative approach consists of monitoring behaviors that are influenced by stress. For instance, Zimmermann et al. [87] proposed monitoring the use of computer mouse and keyboard to capture changes associated with the affective states of users. However, this approach captures indirect information of stress that is only available when the user is performing a specific behavior at the instrumented lo-

cation. Finally, an alternative approach that addresses the previous limitations consists of monitoring the physiological responses associated with stress, such as heart rate, blood volume pulse, skin temperature, pupil dilation or electrodermal activity [6].

The work presented in this thesis considers the last alternative where some physiological signals are monitored. The main idea is to extract objective data to avoid possibles incongruent results in stress scales.

2.4 Stress Recognition Using Physiological Signals

Next question is whether or not stress levels could be recognized accurately from physiological signals using non-intrusive sensors. As was mentioned before, physiological signals are needed in order to extract information about the state of each individual. A novel of the most relevant physiological signals used for detection stress is described below to have a reference of each one, and them, a summary of some experiments is discussed.

Blood Volume Pulse (Blood Volume Presure (BVP))

In stress situations, it is known that there are changes in the number of heart beats due to the fact that pulse is increasing. Also, vasoconstrictions, i.e. the narrowing of the blood vessels resulting from contraction of the muscular wall of the vessels, increases in response to stimulus situation that can provoke stress while decreases in response to relaxation. Therefore, BVP can be an useful signal to measure heart activity. One example of this signal is shwon in Figure 2.2.

This signal offers information about the heart beats and about the relative constrictions of the blood vessels. Calculate the distance between each maximum in seconds is applied in order to measure the heart rate. However, to count vasoconstrictions we have to take into account the shape of the envelope of the signal.

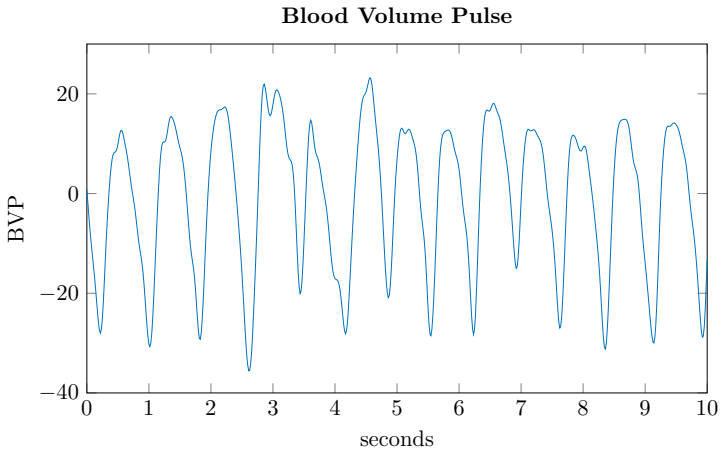


Figure 2.2: Example of a Blood Volume Pressure signal of 10 seconds.

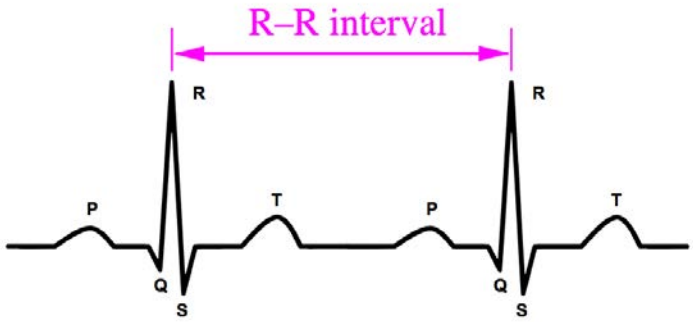


Figure 2.3: ECG signal example and RR interval definition.

Electrocardiogram (ECG)

Electro Cardiogram (ECG) is also used to measure heart activity. It represents graphically the electrical activity of hearts. The electrocardiograph, that is the device which measures the ECG, is able to detect voltage on the surface of the skin when heart beats happen. An example of the signal is shown in Figure 2.3. Inter-Beat (RR) interval is defined as the time elapsing between two consecutive R waves in the electrocardiogram. In order to detect stress, it is necessary to compute other valuable parameter, Heart Rate Variability (HRV). Some researches, for example in [40], show that individuals who have a better stress tolerance, they also have significantly different patterns of Heart Rate Variability in the stress period and also before that.

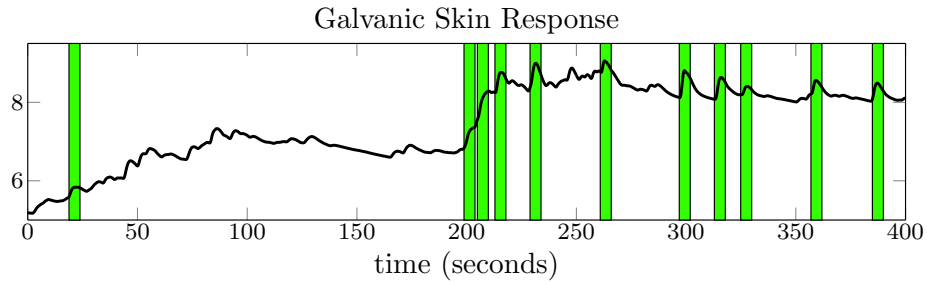


Figure 2.4: Example of Galvanic Skin Response signal of 400 seconds.

Galvanic Skin Response (Galvanic Skin Response (GSR))

Impulses of the sudomotor nerve used to produced and sweat glands react causing sweat diffusion when a stress situation happens. Due to this reason, skin conductance increases and transpiration can be detected by a conductivity sensor. These are the reactions of the Electrodermal Activity (EDA).

The term GSR refers to changes in the electrical properties on the surface of the skin in response to sweat secretions [10]. These secretions are due to an increment in sudomotor innervation, caused by the SNS, that results in changes in the GSR signals as the body responds to different daily circumstances: stress, temperature, anxiety, exertion situations, etc. [76]. Hence, the sympathetic activity can be measured by analyzing the GSR signals, as already shown in[49]. For instance, this relation between SNS acts and the observed skin conductance [79], which determines the sympathetic innervation of the sweat glands [11], has been extensively used in stress detection applications (e.g., see [26, 85, 74, 28]).

GSR suffers changes in the electrical properties of a person's skin caused by an interaction between environment events and the individual's psychological state. However, there exists a big limitation in this signal. In the signal example shown in Figure 2.4, the short stress responses caused by stress are marked in green and the signal is drawn in black. In fact, this example shows that this kind of signals can be useful for stress detection if are processed correctly.

Electromyogram

When the body suffers some stimulus, muscle activation reacts, which supposes an increasing in the current measured in muscles. Therefore, by measuring muscle activity we will be able to detect these current increasings corresponding to stress situation. Apparently, it can be measured in any muscle, but there are some of them that can provide more information, i. e. [57] measure the EMG on the trapeziums muscle giving good results.

As it was explained in the electrocardiogram, in order to measure these reactions that appear when a muscle is contracted we need a set of electrodes. For example in [26], they use three electrodes: two of them located along the axis of the muscle that they want to measure and the third one is the ground located out of this muscle.

As it is described, there is a significant increasing in the current measured that means that the body suffers some stimulus and muscle activation reacts.

Pupil Dilation

Pupil Dilation is wider and the eye movement is faster under a stress situation. Therefore, this is other physiological signal that can be measured in order to know the emotional state of an individual because the Autonomous Nervous System innervates some of the muscles that controls the pupil size.

Pupil Dilation is not used as other biosignals, such as, for instance, skin conductance, because it is more difficult to obtain the acquired signal, but there are several researches where they measured for stress detection [24] [85].

Cortisol Hormone

The conditions of each person can influence at the moment of undergoing activities satisfactorily. Tiredness, family problems, distractions or anxiety are examples of situations where a person's performance can decrease. [51] argues that the measure of cortisol can be an identifier pattern associated with stress as well as in terms of efficiency of individual performance.

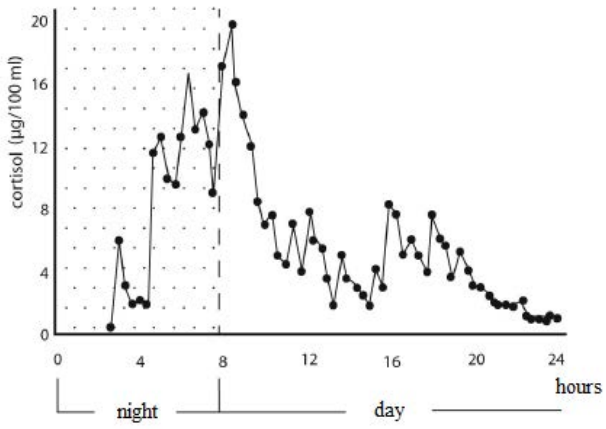


Figure 2.5: Twenty-four hour pattern of cortisol levels [81]

The hormone cortisol is necessary for the functioning of almost every part of the body. Excesses or deficiencies of this hormone also lead to various physical symptoms and disease states. It has been found that the level of salivary cortisol serves as a biochemical marker for post-traumatic psychological distress disorders and other conduct disorders [78]. However, cortisol achievement in the body fluctuate throughout the day with levels, being highest in the morning and lowest in the evening as is presented in Figure 2.5.

Skin Temperature

Skin temperature can change due to the sympathetic innervation of the sweat glands. The main information that can be obtained about the skin temperature is in the transient decreases when the stimulus affects to the individual. Once the acquired signal is measured, the information about when stimulus occurs due to the fact that the slope of the skin temperature signal generally shows a negative trend.

Skin Temperature is not so used because the signal obtained does not provide too much information about the emotional state of the individual. However there exist some researches where this signal is measured and used in order to extract information about it [6].

2.5 Physiological Sensing

A first step towards unobtrusively measuring daily stress consists of the development of tools that can monitor relevant cues of stress without creating additional stress. The current standard approach for measuring heart rate (electrocardiogram) requires sticky gels and uncomfortable electrodes attached to the skin which can be quite cumbersome. Moreover, existing physiological sensors require maintenance (e.g., recharging batteries, replacing electrodes) preventing many people from regularly measuring their vital signs.

One of the least invasive physiological measurement approaches to measure cardiac and respiratory information is Photoplethysmography (PPG), which captures color variations of reflected light from, or transmitted through the skin. Traditional measurements of PPG require a dedicated light source in close contact with the surface of the skin such as the finger [4]. Researchers have also started to consider more wearable approaches which are more appropriate for daily life monitoring. For instance, Kwon et al. [43] attached a smart-phone to the chest and used its accelerometer to monitor heart rate. Similarly, Phan et al. [60] proposed a different method to extract both heart and breathing rates. While moving us one step closer to unobtrusive physiological monitoring during daily life, they mostly considered the chest location where both cardiac and respiratory motions are more prominent. Recently, most of the researchers use several type of smart-wrists to know the information related to hearth rhythm. In the end, some many studies only use the mean heart rate to perform and statistical model so the wave of the electrocardiogram is insignificant.

In the other hand, electrodermal activity signals are more sensitive to recalled a useful signals for signals changes or artifacts. In terms of non intrusive, Shimmer [73] sensor are take placed two electrodes on the palmar surface of the middle and index finger. These sensors could be intrusive, so, some wrist band appear in recent years, such as, Empatica E4 [19] or Q sensor [1] which the bracelet should be placed on the wrist of the non-dominant hand. Wear it snugly, so that it does not move around, but not so snugly that it is uncomfortable.

This thesis systematically evaluates the possibility of measuring heart and electrodermal activity from different wrist with commercially available wearable sensors such as accelerometers and gyroscopes.

This thesis only use non-intrusive sensor to avoid external factors. The sensors used send signals via *Bluetooth* which is a technology standard for exchanging data over short distances. *Bluetooth* is a standard wire-replacement communications protocol primarily designed for low-power consumption, with a short range based on low-cost transceiver microchips in each device. Because the devices use a radio communications system, they do not have to be in visual line of sight of each other. This technique leave behind sensors that must need the acquisition via link because the information extracted could be also acquired via wireless without complications.

The wireless sensors used in this thesis are briefly detailed below:

- Empatica E4 [55] is a compact wristband that allows obtaining data about electrodermal activity, heart rate, R-R interval, device position and angular velocity. This bracelet is shown in Figure 2.6 (a).

- Q sensor [1]. It allows only the measurement of galvanic skin activity through two silver electrodes placed on the base of the device. The quality obtained is better than Empatica E4. This sensor is displayed in Figure 2.6 (b).

- Speech is obtained using a Zoom H1 handheld recorder with a Rode lavalier microphone. Non lossy compressed files (wav), 24-bit quantification and 44 kHz sampling rate are used.

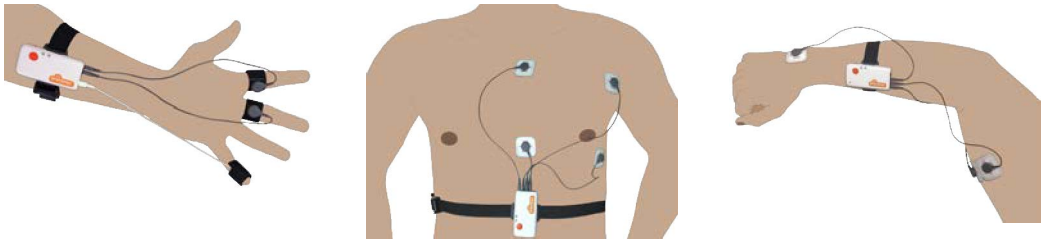
- A Shimmer3 'GSR Unit' sensor was used to capture BVP and GSR signals. On the left hand, the BVP optical pulse sensor was placed on the palmar surface of the pinky finger [72]. For GSR signals two electrodes on the palmar surface of the middle and index fingers were placed [71] as shown in Fig. **??**.

- Another sensor model, Shimmer3 'ExG Unit', was used to capture EMG and ECG signals. In the case of ECG, Fig. **??** shows how the electrodes were positioned on the chest [69].

a: Empatica E4

b: Affectiva Q Sensor

Figure 2.6: Two sensor are shown: (a) Empatica E4 captures BVP and GSR signals and (b) Q sensor acquires only GSR signals.



a: Shimmer used for GSR and BVP signals acquiring electrodes and optical pulse sensing probe

b: Shimmer used for the ECG measurement electrodes

c: Shimmer used for the EMG measurement electrodes

Figure 2.7: Family of Shimmer sensors is displayed.

- In order to capture EMG signals, Fig. **??** shows an example of right arm electrode layout. Two electrodes were placed in parallel with the muscle fibres of the biceps, near the centre of the muscle and the reference electrode an electrically neutral point of the body, as far away as reasonably possible from the muscle being measured [70].

2.6 Stress Recognition In Controlled Environments

Several automatic stress recognition techniques have been explored in the research literature. In most cases, data are collected in the laboratory where variables that may introduce noise are partially controlled or eliminated. Traditional methods

for inducing stress and some complete experiments are described below.

2.7 Elicit Stress Using Controlled Experiments

Stroop Color-Word Interference Test [39]

It is a test in which individuals are presented with lists of color words in matching and non-matching colors and the time they take to read the different words, or the number of errors they make, is recorded. The Stroop effect is the degree of difficulty people have with naming the color of the ink rather than the word itself. There is interference between the color of the ink and the word meaning. This interference occurs no matter how hard you try, which means that it is uncontrollable with the best conscious effort. It implies that at least part of our information processing occurs automatically.

Trier Social Stress Test (TSST) [41]

The TSST is designed to exploit the vulnerability of the stress response to socially evaluative situations. Most current implementations follow a pattern similar to the following:

The period of induced stress lasts approximately 15 minutes, and is divided into 5 minute components. The first 5 minute component is the anticipatory stress phase, during which the judges ask the participant to prepare a 5 minute presentation. The participant is allowed to use paper and pen to organize their presentation, but this paper is then unexpectedly taken away from them when it is time to begin the presentation.

During the 5 minute presentation component, the judges observe the participant without comment. If the participant does not use the entire 5 minutes, they will ask him or her to continue. This goes on until the entire 5 minutes have been used.

The presentation is immediately followed by the mental arithmetic component, during which the participant is asked to count backwards from 1,022 in steps of 13. If a mistake is made, then they must start again from the

beginning. This component lasts for 5 minutes and is followed by a recovery period.

Trier Mental Challenge Test (TMCT)

This test is a selection of mathematical problems. The computer screen displays the actual score un the upper left corner, the mathematical problems in the center and a horizontal bar extending from the left to the right margin on the bottom controlling the time to answers.The time allotted for problem solution varies varies with the difficulty with increasing time for more difficult problems.

After 5 minutes, the computer program stops and only the total scores are displayed. The subjects had to work on 4 trials of arithmetic problems of 5 minutes, this, the stress lasted for 20 min in total.

2.8 Stress Recognition from Physiological Signals

Researchers have explored a variety of classification methods, and techniques to minimize physiological differences across people. Barreto, Zhai & Adjouadi [7], for example, used Support Vector Machines (SVMs) to discriminate between stressful and non-stressful responses (elicited by different versions of the Stroop Test) [39] in a laboratory setting. The SVMs outperformed other classification algorithms, obtaining an accuracy of 90.1%. Various physiological signals were used in the classification, including EDA, BVP, pupil diameter and skin temperature.

In a separate study, Setz et al. [67] used EDA to automatically distinguish between cognitive load and stress elicited by arithmetic computations without and with time pressure and social-evaluative threat, respectively. In this case, Linear Discriminant Analysis (LDA) obtained 82.8% accuracy, outperforming SVMs. Setz et al. [67] found that the average number of EDA peaks, as well as the distribution of their amplitudes, were the most relevant features to the problem. To account for participant variability, distributions were computed for each participant independently.

In another study, Shi et al. [68] discriminated between stressful and non-stressful responses under social, cognitive and physical stressors. They obtained 68% precision (a.k.a., positive predictive value) and 80% recall (a.k.a., sensitivity) using SVMs with EDA, electrocardiogram, respiration and temperature. The problem of participant variability in terms of physiological variance was addressed by subtracting a person specific parameter from the features of each participant. This parameter was estimated as the average feature of all-non-stressful events of the participant.

Finally, another example to comment was to automatically recognize stress in less controlled settings, Healey & Picard [27] monitored electrocardiogram, electromyogram of the trapezius (shoulder), EDA and respiration from people during a real world driving task. They used LDA to automatically discriminate between low (at rest), medium (highways) and high (city) levels of stress with 97% accuracy.

3

Sparse Non-Negative Driver Model

This section presents a new method of feature extraction of Galvanic Skin Response (GSR) signals. It is divided as follows: a signal processing review of GSR methods is discussed, then, the mathematical sparse model is explained, a practical implementation of the model in 100 records is examined and finally some conclusions are considered.

3.1 Signal Processing Review For GSR signals

GSR signals carry information about Sympathetic Nervous System (SNS) activity, but are also influenced by other factors, like temperature changes or sweating due to aerobic exercise [37]. The challenge in analyzing them is thus to develop a method which is able to extract only SNS activity symptoms while avoiding other unrelated components. Indeed, GSR signals, which are usually denoted as $s(t)$,

can be expressed as a sum of two components [10]:

- The *tonic component*, $s_\ell(t)$, or Skin Conductance Lelvel (SCL), which is a slow changing signal. The SCL is related to several non-SNS activity factors but also to the level of attention of the subject, even in the absence of instantaneous stimuli.

- The *phasic component*, $s_p(t)$, also called Skin Conductance Response (SCR), which is the reaction to sporadic SNS stimuli. The SCR, which is superimposed on top of the tonic component, includes higher frequency components and appears only within specific time windows whose length typically lasts from one to five seconds [10].

Furthermore, SCRs can be modeled as the standard linear convolution between a sudomotor SNS innervation, $d_p(t)$, that corresponds to the *non-negative* unknown *sparse* driver that causes the observable skin conductance response, and the response triggered by that driver, $r(t)$. Hence, GSR signals can be finally decomposed as [45],

$$s(t) = s_p(t) + s_\ell(t) = d_p(t) * r(t) + s_\ell(t). \tag{3.1}$$

GSR signal processing models then focus on the estimation of the SCR and SCL components in the absence of other factors.

Alexander et al. [3] were the first to introduce a decomposition algorithm based on the model of Eq. (3.1). Their approach estimates first the SCL contribution, which is subtracted from $s(t)$, and then reconstructs the SCR signal using an iterative inverse filter deconvolution method. However, this method leads to a non-sparse driver that can present negative impulses, which are not physiologically interpretable. Moreover, it is very slow, thus preventing its application for long time registers or on-line signal extraction.

Benedek et al. [9] addressed the decomposition using a different signal model and considering a non-negative deconvolution scheme based on Gauss elimination to avoid negative SCRs. After estimating the SCL and subtracting it from $s(t)$,

this approach obtains an initial estimation of the SCR through a Gauss elimination deconvolution. The negative components of the SCR are then removed by introducing an arbitrary waveform that is fitted by minimizing the error with respect to the observed signal. Unfortunately, this method produces a noisy driver which is not sparse, and the waveforms introduced to force the non-negativity are not physiologically interpretable.

In another work, Benedek et al. [8] proposed an alternative non-negative decomposition based on the model of Eq. (3.1) that shows common ground with the method of [3]. It is based on a spectral division deconvolution with a Gaussian window (after removing the estimated SCL component again), and the SCR detection is performed by searching for the zeros in the first derivative of the driver. This approach yields an individual estimate of the typical SCR shape through optimization, but the estimated driver is still non-sparse and the computation is slow (as shown in the simulations).

Greco et al. [23] proposed a non-negative sparse deconvolution based on cubic B-splines for the SCR component, an Autoregressive–Moving-Average Model (ARMA) model for the sudomotor SNS innervation, and an additive white Gaussian noise term. Although the resulting convex optimization problem can be solved, the solution obtained is still not sparse (values close to zero, but not exactly equal to zero, are obtained) and the interpretability is not improved w.r.t. previous approaches. They introduce the joint estimation of the SCL and SCR components, but this method is still not fast enough for on-line application. Another sparse deconvolution technique has been introduced very recently in [38]. This approach recovers a truly sparse driver and takes into account potential discontinuities in the SCL due to motion artifacts. However, they assume that the length and shape of the response $r(t)$ are known, they do not enforce the separation between SCR events that typically occurs due to physiological reasons, and the resulting algorithm is not directly applicable for on-line extraction of SNS information.

In summary, the sparse nature of the driver $d_p(t)$ that triggers the SCR re-

sponse has not been fully exploited by previous methods. Consequently, the interpretability of the decomposition obtained is limited, since true SCR events are difficult to locate in the driver and artificial signals that have no physiological interpretation are introduced by some methods (e.g., [9]) to reduce the error of the model. Furthermore, none of these approaches is able to provide the real-time results required for on-line operation. We addresses all these issues, developing a novel non-negative sparse deconvolution method (SparsEDA), that introduces the following main contributions:

- Multi-scale analysis that addresses the variable time width of the SCR impulses by using an *overcomplete dictionary* that includes responses, $r(t)$, with different widths and selects the appropriate width for each driver's impulse automatically.

- Exploitation of the sparsity of the SCR component (in order to obtain an interpretable decomposition) by formulating the estimation of $d_p(t)$ as a *sparse inference* problem.

- Fast and efficient solution of the resulting optimization problem, thus allowing for on-line extraction of the SNS information.

- Fully automated implementation of the algorithm in Matlab (released through a free web-based repository) that requires only the selection of two easily interpretable parameters by the user.

Following subsection describes the core of the SparsEDA algorithm: the sparse non-negative deconvolution model used. This includes the description of the discrete-time equivalent of Eq. (3.1) in Section 3.2.1; the multi-scale analysis to account for the variable width of the impulses in Section 3.2.2; the approach used to model the SCL in Section 3.2.3; the optimization problem for the joint estimation of the SCL and SCR components in Section 3.2.4; and the post-processing stage in Section 3.2.5. Then, Section 3.3 addresses three issues that are essential to obtain a robust and efficient implementation of the SparsEDA method: the preprocessing

stage (Section 3.3.1); the continuous mode of operation for on-line signal recovery (Section 3.3.2); and the feature extraction (Section 3.3.3). Finally, Section 3.4 validates the method on 100 signals from 100 different subjects and Section 3.5 provides some concluding remarks.

3.2 Proposed Sparse Model

3.2.1 Discrete-Time Model

The discrete-time model that we consider in the sequel is the following:

$$s[n] = s_\ell[n] + d_p[n] * r[n] + w[n]. \tag{3.2}$$

where $n = 0, \ldots, N-1$ with N denoting the total number of samples available; $s_\ell[n]$, $d_p[n]$ and $r[n]$ are obtained through uniform sampling of $s_p(t)$, $d_p(t)$ and $r(t)$ in Eq. (3.1), with a sampling frequency $f_s = 1/T_s$ Hz; and $w[n]$ is an Additive White Gaussian Noise (AWGN) term that takes into account both measurement noise and discretization error.

Note that (3.2) is not the discrete-time equivalent of (3.1) (as obtained by applying the bilinear transformation [22]), but its sampled version. However, it is a standard model [45, 3, 8, 38].

For finite-length sequences, Eq. (3.2) can be expressed more compactly in matrix form as

$$\vec{s} = \vec{s_\ell} + \vec{R}\vec{d_p} + \vec{w}. \tag{3.3}$$

where $\vec{s_\ell} = [s_\ell[0], \ldots, s_\ell[N-1]]^T$, $\vec{R}$ is an $N \times N$ Toeplitz matrix, $\vec{d_p} = [d_p[0], \ldots, d_p[N-1]]^T$ is an $N \times 1$ sparse non-negative vector, and $\vec{w} = [w[0], \ldots, w[N-1]]^T$ is the noise vector. The vector $\vec{d_p}$ is sparse, since its number of non-null elements (indicated by its L_0 pseudo-norm, $\|\vec{d_p}\|_0$) is small compared to its length, i.e., $\|\vec{d_p}\|_0 = L\ell N$ [18]. Note that a universally accepted threshold to define a vector as sparse does not exist, but we may consider that a vector is sparse when it contains less than 10 % of non-null elements (i.e., $L/N < 0.1$). Since $r[n] = 0$ when $n < 0$

or $n > M-1$ (see Section 3.2.2),

$$\vec{R} = \begin{bmatrix} r[0] & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ r[1] & r[0] & \cdots & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & & & \ddots & & \\ r[M-2] & r[M-3] & \cdots & r[0] & 0 & \cdots & 0 & 0 \\ r[M-1] & r[M-2] & \cdots & r[1] & r[0] & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots & \ddots & & \\ 0 & 0 & \cdots & 0 & 0 & \cdots & r[0] & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & r[1] & r[0] \end{bmatrix}.$$

Our global aim is inferring both $d_p[n]$ and $s_\ell[n]$ jointly when $r[n]$ (and thus $\vec{R}$) is unknown. In order to achieve this goal, we describe first how to approximate $\vec{R}$ in Section 3.2.2 and then how to model $s_\ell[n]$ in Section 3.2.3.

3.2.2 SCR Model: Multi-scale Analysis

We assume that the sudomotor nerve activity can be described by a biexponential function [3] regarding the specific response triggered by the driver:

$$r(t) = e^{-t/\tau_2} - e^{-t/\tau_1}, \qquad \text{for} \quad t \geq 0. \tag{3.4}$$

According to [3], the optimum performance in their experiments is obtained by setting $\tau_2 = 0.75$ and $\tau_1 = 2$. Therefore, we will use these values in the sequel. Another assumption made in [3] is that the duration of a response varies between one and five seconds. Thus, if τ_1 and τ_2 are fixed, in order to construct several waveforms with different time scales we have to use different sampling periods when discretizing $r(t)$. This leads to an overcomplete dictionary, which is a common approach to circumvent the scale problem in sparse inference methods [18]. Using Q different sampling periods, we have

$$\vec{R}_{SCR} = [\vec{R}_1,\ \vec{R}_2,\ \ldots,\ \vec{R}_Q], \tag{3.5}$$

where $\vec{R}_{SCR}$ is an $N \times NQ$ matrix, and the $\vec{R}_q$ $(q = 1, \ldots, Q)$ are a collection of $N \times N$ matrices (the elements of the dictionary) constructed using a predefined set of waveforms $(r_1[n],\ \ldots,\ r_Q[n])$ with different scales. If these waveforms are

carefully chosen, they will be able to provide a very good approximation of the SCR, even if its scale changes over time. The selection of the element used at each time instant is performed by the $NQ \times 1$ extended vector,

$$\vec{d}_{SCR} = [\vec{d}_1^T,\ \vec{d}_2^T,\ \ldots,\ \vec{d}_Q^T]^T, \tag{3.6}$$

which is still a sparse vector with $\|\vec{d}_p\|_0 \leq \|\vec{d}_{SCR}\|_0 \ell NQ$. Hence, the signal model becomes

$$\vec{s} = \vec{s}_\ell + \vec{R}_{SCR}\ \vec{d}_{SCR} + \vec{w}. \tag{3.7}$$

We propose using $Q = 5$ waveforms,

$$r_i[n] = e^{-nT_s/\tau_{2i}} - e^{-nT_s/\tau_{1i}},$$

for $n = 0, \ldots, M-1$ and $i = 1, \ldots, 5$, with $\tau_{1i} = k_i\tau_1$, $\tau_{2i} = k_i\tau_2$, $k_i \in \{\frac{1}{2},\ \frac{3}{4},\ 1,\ \frac{5}{4},\ \frac{3}{2}\}$, and $M = 10/T_s$ (i.e., M is the number of samples obtained in a time interval of 10 seconds using a sampling period T_s). Therefore, the SCR matrix finally becomes:

$$\vec{R}_{SCR} = [\vec{R}_{T_1},\ \vec{R}_{T_2},\ \vec{R}_{T_3},\ \vec{R}_{T_4},\ \vec{R}_{T_5}], \tag{3.8}$$

where $\vec{R}_{T_i}$ is constructed using $r_i[n]$ for $i = 1, \ldots, 5$. Note that this matrix allows us to cater both for short time-scale processes (e.g., 1-2 seconds) and long time-constant processes (e.g., 3-5 seconds). Furthermore, since we construct the periods $T_1, \ldots, T_5$ as a function of T_s, we are also able to deal with any sampling period automatically. Fig. 3.1 shows the shape of the five different SCR waveforms used to construct $\vec{R}_{SCR}$. Note that all of them correspond to an SNS occurring at the same time instant $t = 0$ (i.e., without any delay), although their effective lengths and the positions of the peaks of the SCR responses are different. The peak of the SCR response is located $\frac{6\ln(8/3)}{5}k_i$ seconds after the SCR event that triggered the response. For the $k_i \in \{\frac{1}{2},\ \frac{3}{4},\ 1,\ \frac{5}{4},\ \frac{3}{2}\}$ used in Fig. 3.1 this corresponds approximately to $\{0.5885,\ 0.8827,\ 1.1770,\ 1.4712,\ 1.7655\}$ seconds, respectively.

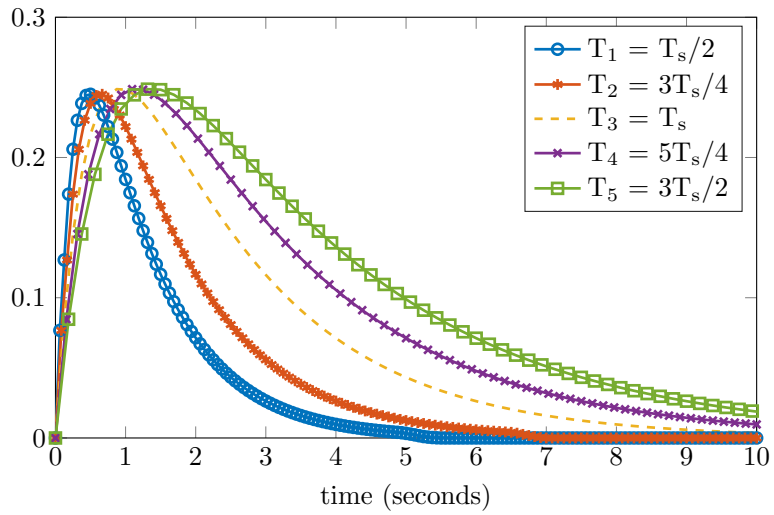


Figure 3.1: Multi-Scale SCR waveforms ($r_1[n]$, ..., $r_5[n]$), obtained discretizing $r(t)$ in Eq. (3.4) using five different periods: $T_1 = T_s/2$, $T_2 = 3T_s/4$, $T_3 = T_s$, $T_4 = 5T_s/4$, and $T_5 = 3T_s/2$. All of them correspond to an SNS occurring at the same time instant $t = 0$.

3.2.3 SCL Model: Taylor Series Expansion

In order to approximate the SCL in a simple and efficient way, we propose to use a first order Taylor series expansion:

$$\vec{s}_\ell = [\vec{1}/||\vec{1}||_2 \;\; \boldsymbol{\ell}/||\boldsymbol{\ell}||_2 \;\; -\boldsymbol{\ell}/||\boldsymbol{\ell}||_2]\, \vec{d}_{SCL} = \vec{R}_{SCL}\vec{d}_{SCL}, \tag{3.9}$$

where $\mathbf{1} = [1, \ldots, 1]^T$ and $\boldsymbol{\ell} = [0, 1, \ldots, N-1]^T$ are constant and linearly increasing $N \times 1$ column vectors, respectively; $\|\cdot\|_2$ denotes the L_2 norm, so $||\vec{1}||_2 = \sqrt{N}$ for instance; $\vec{R}_{SCL}$ is the $N \times 3$ matrix built using those column vectors; and $\vec{d}_{SCL}$ is a non-negative coefficients vector.

3.2.4 Joint SCL and SCR estimation

The standard approaches to GSR signal analysis in the literature are based on the sequential extraction of $s_\ell(t)$ and $s_p(t)$. On the one hand, $s_\ell(t)$ can be estimated as an average of $s(t)$ over short intervals (10 to 100 seconds [3, 9]), and then subtracted from $s(t)$ in order to obtain $s_p(t)$. Alternatively, $s_p(t)$ can be obtained from $s(t)$ through a high pass filter [21], and then, if desired, $s_\ell(t)$ can be obtained

again by subtraction. Either way, as pointed out in [8], this sequential extraction tends to underestimate the SCR component.

Here, it is proposed to estimate both the SCL and the SCR components simultaneously, exploiting the fact that $\vec{R}_{SCL}$ and $\vec{R}_{SCR}$ can be combined into a single $N \times (NQ + 3)$ matrix: $\vec{R}_T = [\vec{R}_{SCL}\ \vec{R}_{SCR}]$. Hence, can be rewritten Eq. (3.7) more compactly as

$$\mathbf{s} = \vec{R}_T \vec{d}_T + \vec{w}, \tag{3.10}$$

where $\vec{d}_T = [\vec{d}_{SCL}^T\ \vec{d}_{SCR}^T]^T \in \mathbb{R}^{NQ+3}$. This is the joint SCL-SCR model that we use in the sequel to perform the joint estimation of the SCL and SCR components.

The proposed solution is based on the following key characteristics of the skin conductance response [79, 48]:

- The driver, $d_p(t)$, represents sudomotor nerves activations, and thus it corresponds either to non-negative deflections (during states of activity) or remains equal to zero otherwise.
- A single impulse corresponds to each SNS act [8] and triggers an SCR response that typically lasts from 1 to 5 seconds [8].
- The sudomotor impulses arrive as discrete and separate (i.e., non-overlapping) events, implying that $\vec{d}_{SCR}$ should be sparse in time.

These facts lead us to consider a *non-negative sparse driver*, with the constraint that $d_T(i) \geq 0$ for $1 \leq i \leq N$, which is characterized by a zero baseline and occasional (i.e., sparse) discrete positive impulses with a compact support. Hence, the estimation of $\vec{d}_T$ can be formulated as the following constrained optimization problem:

$$\hat{\vec{d}}_T = \underset{\vec{d}_T}{\arg\min} \quad \|\vec{s} - \vec{R}_T \vec{d}_T\|_2^2 \tag{3.11a}$$

$$\text{subject to} \quad d_T(i) \geq 0\ \forall i \tag{3.11b}$$

$$\|\vec{d}_T\|_0 \ell NQ \tag{3.11c}$$

Eq. (3.11a) corresponds to the minimization of the mean squared error Mean Square Error (MSE) between the available signal and the reconstructed one, Eq.

(3.11b) is the non-negativity constraint, and Eq. (3.11c) is the sparsity constraint. Since the L_0 pseudo-norm is untractable from a mathematical point of view, the sparsity constraint in Eq. (3.11c) can be imposed using an L_1 regularization term, that leads to the following non-negative version of the Least Absolute Shrinkage and Selection Operator (LASSO) [77]:

$$\hat{\vec{d}}_T = \underset{\vec{d}}{\arg\min} \quad \|\vec{s} - \vec{R}_T \vec{d}_T\|_2^2 + \lambda \|\vec{d}_T\|_1 \tag{3.12a}$$

$$\text{subject to} \quad d_T(i) \geq 0 \; \forall i \tag{3.12b}$$

which can be solved efficiently using the Least Angle Regression (LARS) algorithm [17].

Note that the regularization parameter, λ, controls the amount of sparsity in the reconstruction: the larger the value of λ the sparser the signal obtained. However, λ has no clear interpretation in terms of GSR signals, and thus it can be difficult for the user to find an appropriate value for it. Instead, since LARS is a greedy algorithm, we propose to replace the choice of λ by the selection of a stopping rule that can be set much more easily by users without a detailed technical knowledge of the algorithm. For instance, we stop the optimization when the residual of the GSR reconstruction is less than a pre-defined value (i.e., $\|\vec{s} - \vec{R}_T \hat{\vec{d}}_T\|_2^2 \leq \epsilon$) or a maximum number of iterations $K_{\max}$ have been performed. Furthermore, since a final post-processing stage is performed (see Section 3.2.5) to remove redundant impulses, the user can simply set values of ϵ and $K_{\max}$ which ensure that all the potential impulses have been discovered by LARS (e.g., in the simulations we have used $\epsilon = 10^{-4}$ and $K_{\max} = 40$), and then control the degree of sparsity during the post-processing.

3.2.5 Post-Processing

The goal of the post-processing stage is to obtain a driver signal, $d_p(t)$, which is as sparse as possible. In order to achieve this goal, we propose to apply a greedy algorithm that eliminates weak impulses which are too close to stronger impulses, following a similar approach to [56, 46]. These impulses are included by LARS

in order to decrease the MSE, but are useless from a physiological point of view and hinder the interpretability of the recovered signals. In summary, the proposed approach is the following:

1. Initialize the set of accepted impulses as the empty set: $\mathcal{A} = \emptyset$.

2. Sort the elements of $\vec{d}_{SCR}$ in descending order according to their L_1 norm. Hence, the resulting ordered vector, $\vec{d}_{ord}$, fulfills that

$$\|\vec{d}_{ord}(1)\|_1 \geq \|\vec{d}_{ord}(2)\|_1 \geq \cdots \geq \|\vec{d}_{ord}(N)\|_1.$$

   Note that, since $\vec{d}_{SCR}$ is sparse, only the first $L_{SCR}\ell NQ$ elements of $\vec{d}_{ord}$ are different from zero.

3. For $i = 1, \ldots, L_{SCR}$:

   (a) If the location of $\vec{d}_{ord}(i)$ is not within $T_{\min}$ seconds (i.e., $N_{\min} = T_{\min} f_s$ samples) of an existing impulse in $\mathcal{A}$, add it to the set of accepted impulses.

   (b) Otherwise, discard it.

4. Discard the accepted impulses whose L_0 norm lies below the following threshold:

$$\gamma = \rho \max \|\vec{d}_{SCR}\|_1,$$

   where $0 < \rho < 1$ is a user specified parameter that can be used to control the final sparsity of the solution obtained.

The minimum distance constraint enforced by the previous algorithm could also be included within the optimization problem, as shown by the Cross-Products LASSO algorithm [47]. However, this leads to a substantial increase in computational cost (since the resulting problem is not convex anymore), that would prevent the on-line implementation that we are seeking here.

3.3 Practical Implementation

3.3.1 Preprocessing

Continuous and unobtrusive measurement of GSR using wearable devices makes the signal collected vulnerable to several types of noise and artifacts. Artifacts can be generated from electronic noise or variation in the contact between the skin and the recording electrode caused by pressure, excessive movement or adjustment of the device. If these artifacts remain in the signal when it is analyzed, they can easily be misinterpreted and skew the analysis; for example, they may be easily mistaken for an SCR [76].

In order to remove these artifacts, Sara Taylor et al. [76] developed a machine learning algorithm to detect automatically Electrodermal Activity (EDA) artifacts that will be used before applying our novel SparsEDA algorithm. This method takes segments of 5 seconds and classifies them either as artifacts or as valid GSR signals. In those slots that are classified as artifacts, the signal is replaced by a polynomial regression of order 1 between the last non-artifact point in previous slots and the first non-artifact sample in subsequent slots, and no further processing is performed.

Besides the artifact removal stage, the complete signal is resampled to 8 Hz if the sampling frequency, f_s, is higher than 8 Hz. This resampling does not have any influence on the signal's quality, since SCR waveforms can be perfectly represented using $f_s = 8$ Hz (or even $f_s = 4$ Hz), but allows us to reduce notably the computation time.

3.3.2 Continuous-mode Operation

The final goal of the SparsEDA method is being able to extract EDA-related features continuously using shorter signal sets of $N = W + M + 1 < L$ samples, where M is the length of the waveforms used to construct the SCR dictionary and L is the total number of samples available of the GSR signal. This continuous-mode operation allows us to deal with large signals of arbitrary length (since the approach

$$\vec{R}'_{SCL} = \begin{bmatrix} \bar{\ell}_1(1) & -\bar{\ell}_1(1) & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \bar{\ell}_1(\frac{W}{3}) & -\bar{\ell}_1(\frac{W}{3}) & 0 & 0 & 0 & 0 \\ \bar{\ell}_1(\frac{W}{3}+1) & -\bar{\ell}_1(\frac{W}{3}+1) & \bar{\ell}_2(1) & -\bar{\ell}_2(1) & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \bar{\ell}_1(\frac{2W}{3}) & -\bar{\ell}_1(\frac{2W}{3}) & \bar{\ell}_2(\frac{W}{3}) & -\bar{\ell}_2(\frac{W}{3}) & 0 & 0 \\ \bar{\ell}_1(\frac{2W}{3}+1) & -\bar{\ell}_1(\frac{2W}{3}+1) & \bar{\ell}_2(\frac{W}{3}+1) & -\bar{\ell}_2(\frac{W}{3}+1) & \bar{\ell}_3(1) & -\bar{\ell}_3(1) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \bar{\ell}_1(W) & -\bar{\ell}_1(W) & \bar{\ell}_2(\frac{2W}{3}) & -\bar{\ell}_2(\frac{2W}{3}) & \bar{\ell}_3(\frac{W}{3}) & -\bar{\ell}_3(\frac{W}{3}) \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{3.13}$$

is applied on segments of fixed length) and even to process signals on-line (i.e., as they are being acquired) in real-time or quasi-real-time. The SparsEDA algorithm incurs in a fixed processing delay equal to the time shift used (W or $W/3$), but afterwards it is able to return the SCL and SCR values almost instantaneously. In order to describe this mode of operation, let us set $M = 10/T_s$ (i.e., the duration of the SCR waveforms is 10 seconds), $W = 60/T_s$, and refer to the current slot being processed (whose duration is 70 seconds) as the *active set*.

First of all, a *naive implementation* of the SparsEDA algorithm for large/on-line signals is straightforward:

1. Extend the signal by adding M samples before the first one and W samples after the last one. These samples are required to ensure that the whole signal is properly processed and can take arbitrary values. For instance, we simply replicate the first and last samples in the signal M and W times, respectively, to perform this extension.

2. Set the first N samples of the extended signal as the active set.

3. Apply the SparsEDA method described in Section 3.2 on the current active

set, obtaining an estimate of its SCL and SCR components for the first W samples.

4. Keep shifting the active set by W samples and repeating the previous stage until all the samples in the original GSR signal have been processed.

5. Discard the SCL and SCR components corresponding to the samples added at the beginning and the end of the signal.

However, this naive approach does not ensure the continuity of the signal among the different segments for the SCL component and can lead to high MSE values. In order to promote the continuity, we propose to divide the active set into three continuous sets, composed of $\frac{1}{3}W$ samples each one, and to use a modified $\vec{R}_{SCL}$ matrix, $\vec{R}'_{SCL}$. This matrix, shown in Eq. (3.13), is composed of W rows (i.e., the length of the active set) and six columns that correspond, respectively, to the positive and negative versions of the SCL model for each of the three segments of the active set. In Eq. (3.13), the $\boldsymbol{\ell}_i$ are linearly increasing column vectors ($\boldsymbol{\ell}_1$ from 0 to 1 with a length $L_1 = W$, $\boldsymbol{\ell}_2$ from 0 to 2/3 with a length $L_2 = 2W/3$, and $\boldsymbol{\ell}_3$ from 0 to 2/3 with a length $L_3 = W/3$); $\bar{\boldsymbol{\ell}}_i = \boldsymbol{\ell}_i / \|\boldsymbol{\ell}_1\|_2$ are the normalized vectors (w.r.t. the L_2 norm of $\boldsymbol{\ell}_1$); and $\bar{\ell}_i(k)$ denotes the k-th element ($k = 1, 2, \ldots, L_i$) of $\bar{\boldsymbol{\ell}}_i$. This construction of the matrix enforces the continuity among those three segments, since a discontinuity results in an increase in the MSE. Hence, the *continuous-mode operation* SparsEDA algorithm proposed is finally:

1. Extend the signal by adding M samples before the first one and W samples after the last one.

2. Set the first N samples of the extended signal as the active set.

3. Subtract the value of the first sample, and apply the SparsEDA method described in Section 3.2 on the current active set, using a matrix $\vec{R}'_T = [\vec{R}'_{SCL}\ \vec{R}_{SCR}]$, with $\vec{R}'_{SCL}$ given by (3.13) and $\vec{R}_{SCR}$ given by (3.8), thus obtaining an estimate of the SCL and SCR components for the first $W/3$

samples. Note that (3.13) does not contain the offset term for the Taylor series expansion used to model the SCL. Hence, we subtract the first sample of the active set to ensure that the line used to model the SCL starts at 0. Add back the value of the subtracted sample to obtain the final solution.

4. Keep shifting the active set by $W/3$ samples and repeating the previous stage until all the samples in the original GSR signal have been processed.

5. Discard the SCL and SCR components corresponding to the samples added at the beginning and the end of the signal.

Note that the difference w.r.t. the naive implementation lies in steps 3 and 4 (using a modified matrix R'_T and shifting the active set by $W/3$ samples instead of W).

The proposed SparsEDA algorithm has been made freely available (altogether with the results presented in Section 3.4, which are provided in a `.mat` file) through a well-known web-based code repository (`https://github.com/fhernandogallego/sparsEDA`) and can also be found in the first author's web-page (`http://www.tsc.uc3m.es/~fhernando/Research.html`). The code, which has been developed in MATLAB, is based on the implementation of the LASSO algorithm provided by [80].

3.3.3 Feature Extraction

The SparsEDA algorithm allows us to extract two GSR components (features) that should be interesting for SNS studies:

- **Tonic component (SCL):** The slope is related to the level of attention of the subject. If the patient is concentrated and/or involved in a task, an increasing slope should be observed [29]. Otherwise, the SCL slope should either decrease or remain constant.

- **Phasic component (SCR):** This is the indicator of SNS reactions. The proposed method allows us to extract both their locations and durations. Furthermore, the sparsity of the resulting driver enhances its interpretability, whereas the post-processing stage allows us to avoid false alarms.

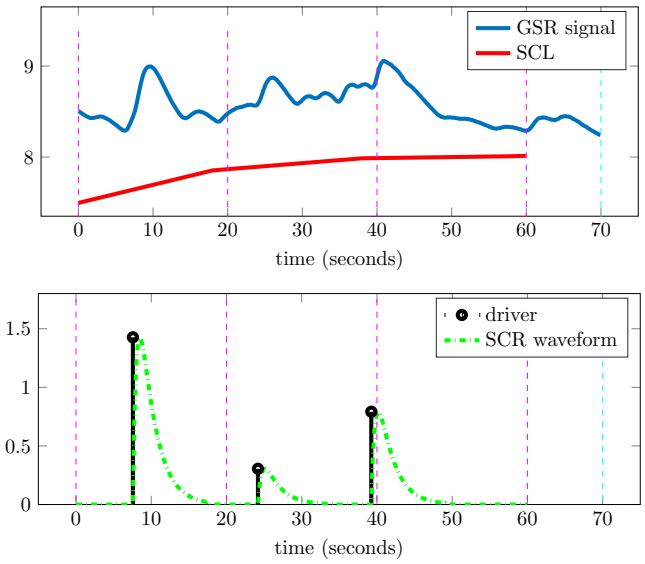


Figure 3.2: Signal extraction example using SparsEDA. **(Top)** A GSR signal (blue) of duration equal to 70 seconds (i.e., length $L = 70/T_s$ samples), and the SCL signal (red) obtained by SparsEDA. **(Bottom)** Sparse driver obtained (black) and the SCR waveform for each mark (dash-dotted green).

An example of the application of the SparsEDA method is shown in Fig. 3.2. The input is a GSR signal of length $N = W + L + 1$ (in blue), and the results (features) are a sparse driver signal of length W (displayed as black impulses), the corresponding SCR component of length W (in green), and the low-frequency SCL components for each length $\frac{1}{3}W$ set (in red) with their associated slope. Purple vertical lines separate each $\frac{1}{3}W$ set, whereas the cyan line indicates the end of the signal.

3.4 Results

First of all, we show the qualitative behavior of SparsEDA on a record acquired at $f_s = 16$ Hz with length equal to 400 seconds that is freely available at [44]. Fig. 3.3 compares the SCL and SCR components extracted by the novel SparsEDA algorithm (without decreasing the sampling rate) and two alternative approaches: the Continuous Decomposition Analysis (CDA) technique introduced in [8] (CDA

Ledalab), and the convex approach proposed in [23] (cvxEDA). Regarding the SCL component, it can be seen that SparsEDA retrieves an SCL component which is very similar to the one returned by cvxEDA, and both of them lie below the SCL component obtained using CDA Ledalab. However, the main difference can be appreciated in the SCR component. On the one hand, the SCR signal returned by CDA Ledalab is not at all sparse and the driver's impulses are very difficult to locate. On the other hand, although cvxEDA returns a sparser SCR component, it still contains too many activations to be useful for the task of locating the driver's impulses and counting their number. Finally, SparsEDA returns a truly sparse SCR signal (with a degree of sparsity that can be easily controlled by the user, as described in Section 3.2) that contains the main impulses of the driver and which is much more interpretable than the other two from a physiological point of view.

Then, in order to show the versatility of the proposed approach, we have compared the performance of the novel SparsEDA method with the same two algorithms as before: CDA Ledalab [8] and cvxEDA [23]. The simulations have been performed on a database composed of 100 GSR signals from 100 different patients acquired at several sampling rates from 4 Hz to 128 Hz, with different sensors and a wide range of signal lengths. 50 signals were recorded within the ES3 project [2], using Medicom MTD sensors [54] and following the procedure described in [30]. The rest were recorded in our laboratory using: the Empatica E4 sensor [19] (5 signals), Microsoft's Band 2 [55] (5 signals), Qsensor [1] (35 signals), and Shimmer3 [73] (5 signals). Regarding our SparsEDA algorithm, we have used the same parameters as before: $\epsilon = 10^{-4}$, $K_{\max} = 40$ iterations, $N_{\min} = \frac{5}{4} f_s$ samples, and $\rho = 0.025$. All the signals have been preprocessed using the web-based method proposed in [76], segmenting them into slots of 5 seconds that correspond either to a valid GSR signal or artifact/noise, and replacing the slots that were labelled as artifacts by a linear regression. After this preprocessing stage, each signal was processed using the three aforementioned algorithms, extracting the SCR and SCL components.

Table 5.1 summarizes the characteristics of the six groups of data used in the simulations (number of signals, sampling rate, type of sensor used and duration

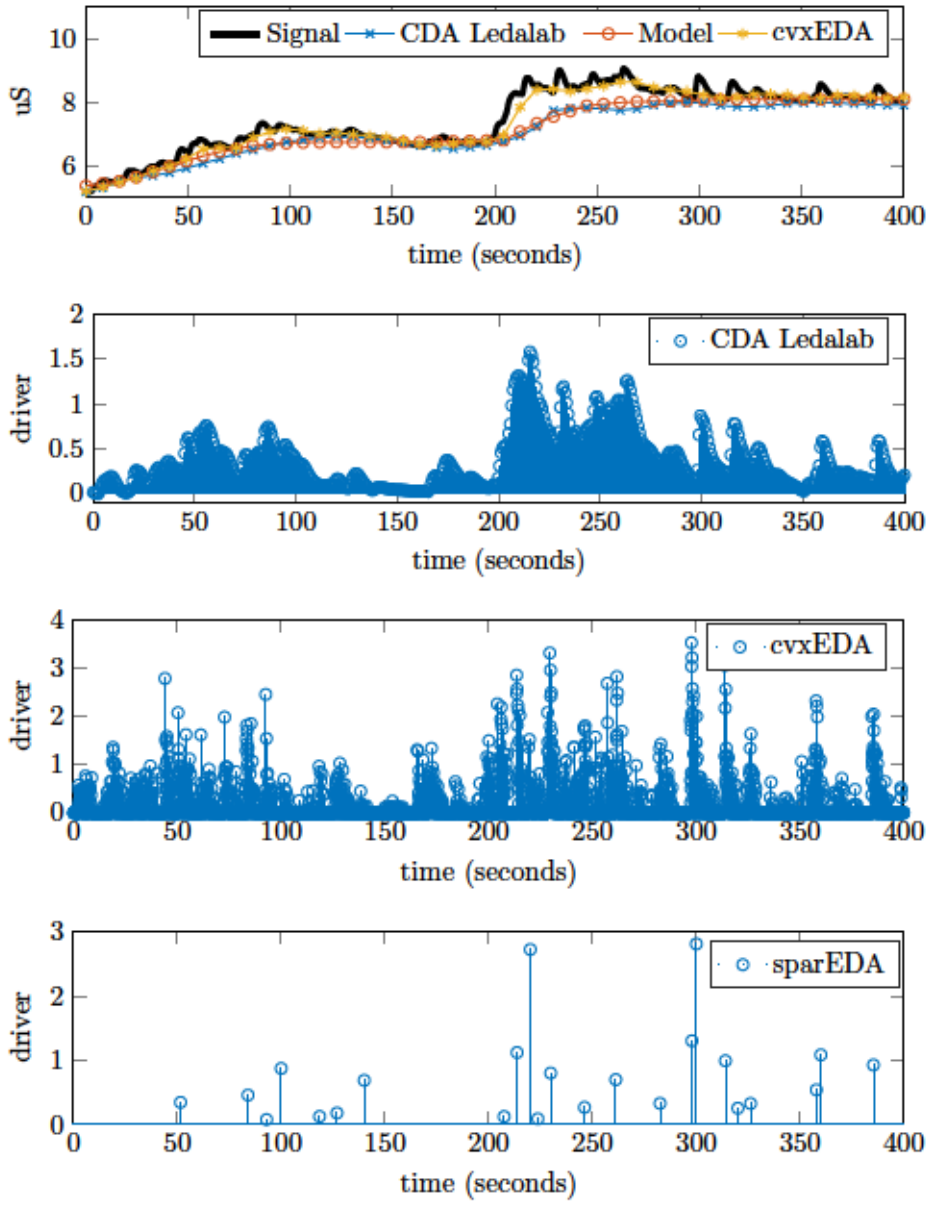


Figure 3.3: Signal processing for GSR signal feature extraction using a sampling rate $f_s = 16$ Hz. (a) Black line representing the GSR signal and several lines with markers showing different SCL approximations: CDA Ledalab (blue crosses), SparsEDA (red circles) and cvxEDA (yellow stars). (b) CDA Ledalab driver obtained using 4 iterations of the optimization tools. (c) Driver obtained using cvxEDA. (d) Driver obtained using SparsEDA with $\epsilon = 10^{-4}$, $K_{\max} = 40$ iterations, $N_{\min} = \frac{5}{4} f_s = 20$ samples, and $\rho = 0.025$.

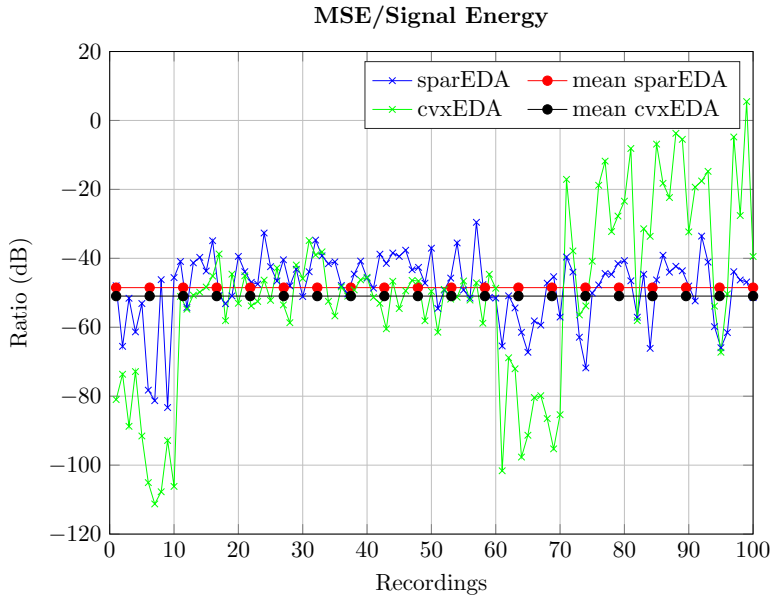


Figure 3.4: Ratio (dB) between the MSE and the signal energy for the 100 signals in the database and two of the methods compared: SparsEDA and cvxEDA. Blue and green lines show the ratio of SparsEDA and cvxEDA for each particular signal in the database, whereas the horizontal red and black lines display their average values.

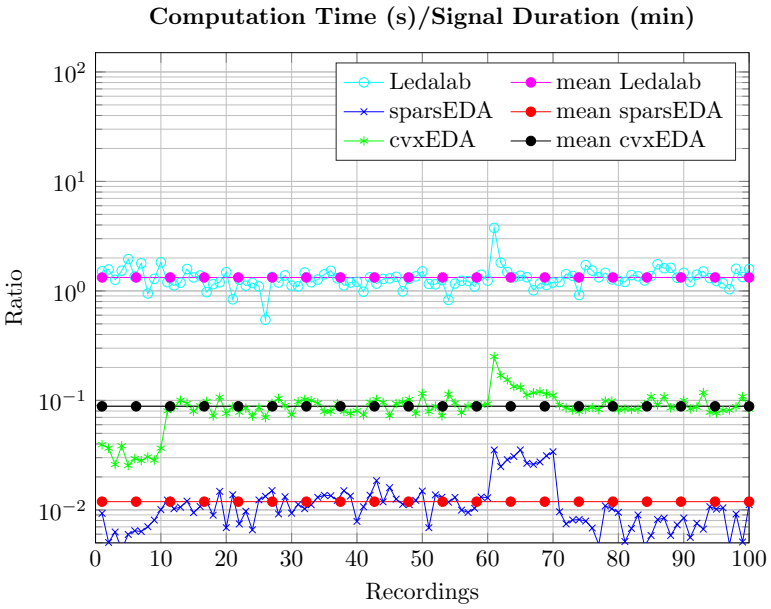


Figure 3.5: Ratio between the computation time (in seconds) and the signal's duration (in minutes) for the 100 signals in the database and the three methods compared: SparsEDA, CDA Ledalab and cvxEDA. Cyan, blue and green lines show the ratio of SparsEDA, CDA Ledalab and cvxEDA for each signal in the database, whereas the horizontal purple, red and black lines display their average values.

[mean and standard deviation (std.)]), altogether with the results obtained: the relative MSE for SparsEDA (mean and std.) and the computation time for the three algorithms tested (mean and std.). Note that sampling rates larger than 8 Hz has been re-sampled to 8 Hz for all of the methods. A detailed analysis of the results obtained is provided in the sequel:

- Fig. 3.4 displays the MSE obtained by SparsEDA and cvxEDA for each record. Let us remark that the main goal of SparsEDA is obtaining an interpretable SCL/SCR decomposition in a small amount of time (i.e., attaining a good balance between computation time and MSE), not fitting the SCL and SCR components in order to obtain a zero remainder. However, very good values of MSE are attained: from a minimum value of -83.29 dB to a maximum value of -29.55 dB, with an average MSE equal to -48.50 dB. Although cvxEDA attains lower MSE values (from -111.26 dB to 5.47 dB, with -50.95 dB on average), this is hindered by the much lower interpretability of the resulting SCR signal (as shown in Fig. 3.3 and discussed below).

- The MSE obtained by CDA Ledalab is not included in Fig. 3.4 because it is fitted to zero by using a non-sparse SCR component (which is not physiologically interpretable) that contains all the elements of the signal that do not belong to the SCL.

- The computation time is summarized for each method in Table 3.1, showing that SparsEDA is always faster than cvxEDA, and both of them are much faster than CDA Ledalab. Fig. 3.5 delves deeper into this important feature, showing the ratio between the computation time (in seconds) and the duration of the signal (in minutes). The average values of the ratios, which are also displayed, are 1.327 for CDA Ledalab, 0.088 for cvxEDA and 0.012 for SparsEDA (i.e., on average SparsEDA is 7.42 times faster than cvxEDA and 111.58 times faster than CDA Ledalab).

- The level of sparsity in the SCR component is directly related to the interpretability of the decomposition obtained: the sparser the signal the more in-

terpretable from a physiological point of view. In this sense, while SparsEDA attains an average sparsity of 0.287% (percentage of non-zero values out of the total number of samples), the other two methods are non-sparse, always returning 100% non-zero values (cvxEDA contains many samples with small but non-null values).

- Confidence intervals and p-values comparing SparsEDA with CDA Ledalab and cvxEDA (both in terms of relative MSE and computation time) have been computed and are displayed in Table 3.2. These values have been computed through the t-test (using Matlab's `ttest2` function) on the hypothesis that the mean of the two distributions of the results (i.e., those of SparsEDA and the compared algorithm) are equal. From the p-values obtained (well below the 0.05 threshold typically used) we can see that both the differences in relative computation time and MSE are statistically significant. However, note that the loss in MSE attained by SparsEDA is small with respect to the decrease in computation time, as shown in Figs. 3.4 and 3.5 and evidenced by the p-values in Table 3.2.

#	Sampling Rate	Sensor	Duration (minutes)		MSE/SE (dB)		Computation Time (seconds) CDA Ledalab		SparsEDA		cvxEDA	
			Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
1-5	4 Hz	Empatica E4	50.00	00.00	-55.90	7.27	78.29	12.36	0.31	0.09	1.65	0.34
6-10	5 Hz	Microsoft Band 2	33.99	19.50	-66.92	19.28	52.51	41.65	0.28	0.23	1.09	0.78
11-60	8 Hz	Medicom MTD	39.43	05.47	-44.03	5.97	48.53	10.37	0.46	0.11	3.50	0.72
61-65	16 Hz	Qsensor	32.62	22.00	-59.88	7.07	41.15	53.46	0.29	0.24	2.73	2.71
66-70	128 Hz	Shimmer 3	9.90	00.00	-53.42	6.62	11.56	1.14	0.14	0.03	1.13	0.03
71-100	8 Hz	Qsensor	9.89	00.00	-48.92	9.20	72.80	10.13	0.41	0.10	4.70	0.55

Table 3.1: Summary of the characteristics of the 100 GSR records used in the simulations and the results obtained. They are separated into six groups according to the sampling rate and the sensor. The duration (mean and standard deviation [std.]) of each group are also shown. Relative MSE (MSE/Signal Energy (SE)) is only shown for SparsEDA. Computation time (mean and std.) is shown for the three algorithms compared: SparsEDA, CDA Ledalab and cvxEDA.

#	Sampling Rate	MSE/SE SparsEDA vs cvxEDA			CT/SD SparsEDA vs CDA Ledalab			CT/SD SparsEDA vs cvxEDA		
		p-values	conf. int.		p-values	conf int.		p-values	conf. int.	
			lower	upper		lower	upper		lower	upper
1-5	4 Hz	9.09 e-04	14.09	37.20	6.21 e-07	-1.81	-1.32	2.96 e-05	-0.03	-0.020
6-10	5 Hz	3.40 e-03	16.54	58.84	2.39 e-09	-1.81	-1.05	8.26 e-07	-0.027	-0.019
11-60	8 Hz	8.67 e-06	3.16	7.79	6.77 e-68	-1.25	-1.15	4.01 e-68	-0.080	-0.073
61-65	16 Hz	7.30 e-03	9.35	43.49	3.10 e-10	-2.98	-0.86	2.50 e-04	-0.18	-0.090
66-70	128 Hz	4.81 e-05	22.68	41.41	1.90 e-08	-1.25	-1.02	2.07 e-10	-0.091	-0.081
71-100	8 Hz	6.58 e-07	-28.89	-13.63	-2.21 e-43	-1.43	-1.30	-3.90 e-45	-0.085	-0.077

Table 3.2: Statistical comparative between SparsEDAand cvxEDAin terms of relative MSE/SE and SparsEDAversus cvxEDAand CDA Ledalab in terms of CT/SD. P-values and confidence intervals (lower and upper limits) are displayed to show the significant differences. MSE/SE = MSE/Signal Energy. CT/SD = Computation Time (s)/Signal Duration (min).

3.5 Conclusions

We have developed a novel feature extraction method for GSR signals: the SparsEDA algorithm. The main contributions of SparsEDA are the joint estimation of the SCL and SCR components, the multi-scale analysis using an over-complete dictionary, the retrieval of a sparse driver for the SCR component, and the efficient implementation of a fully automated, continuous-mode operation algorithm for on-line processing. The proposed approach has been tested on a database of 100 GSR records from 100 different patients acquired using different sensors and sampling rates, confirming its good performance in terms of relative MSE (-48.50 dB) and computation time (1–2 orders of magnitude lower than other existing algorithms like CDA Ledalab and cvxEDA). Furthermore, the interpretability of the SCR component extracted is enhanced w.r.t. previous approaches, thanks to the sparsity of the driver (unlike both CDA Ledalab and cvxEDA, which return non-sparse drivers) and the lack of artificial non-interpretable signals introduced to minimize the MSE (as done by CDA Ledalab). In summary, SparsEDA confirms the feasibility of developing a fast and fully automated method for extracting the GSR components from large EDA records. One possible future line would be embedding this software within a wearable sensor or a real-time mobile application to detect SNS symptoms using driver activations. Another potential application would be integrating it within medical software to detect stress reactions while an patient is performing some activity.

4

Stress Feature Extraction

This chapter presents a novel of the physiological features used in this thesis for stress assessment. Each subsection focus on a human activity: heart rhythm, sweating, speech and hormone analysis.

While research on automated stress recognition has taken many different forms, the systems that have been proposed in the engineering literature typically contain two principle components: 1) a sensor-based architecture that records relevant features and 2) a software-based system that makes predictions about an individual's current stress level. The sensing modalities can take many forms, including audio and visual modalities, but biosensors provide the most direct access into the physiological changes that accompany stress.

The criteria selection for each feature is based on two assumptions: there are features used in the past for another researchers and the acquisition should be

non-intrusive. Some examples of past feature extraction models are: D. Wu et al. [83] captured physiological responses from: Galvanic Skin Response (GSR), respiration, electroencephalogram and electrocardiogram for identification and classification into several stress levels. Another case was Zhai and Barreto [86], who acquired different affective features: GSR, Blood Volume Presure (BVP) and pupil diameter to differentiate states (relax or mental stress) in computers user. Speech also plays a role in detecting stress as can be depicted form [16] where stress is estimated using pitch and speech energy features and [58] where Log Frequency Power Coefficients (LFPC) are used to detect stress and emotion.

The minimum length in time of the experiments proposed in the following section is 30 minutes. As the changes of the Autonomous Nervous System (ANS) are provoked slowly, we accept to extract the mean of the features each minute (60 seconds) to reduce the redundancy in theses feature extraction algorithm.

4.1 Heart Rhythm Activity

The most common feature is Heart Rate Variability (HRV) which analyses the physiological phenomenon of the oscillation in the interval between consecutive heart beats [15].

The proposed features can be divided in time or frequency domain. On one hand, the following features are computed in the time domain:

- Standard Deviation of NN Intervals (SDNN). Computed following the expression:

$$\text{SDNN}_i = \sqrt{\frac{1}{n-1}\sum_{j=1}^{n}(RR_j - \overline{RR})^2}$$

  where RR_J is the inter-beat (RR) interval, $\overline{RR}$ is the mean of the RR over the window, and n is the number of RR points in the window (60 in this case).

- Root Mean Square of Successive Differences (RMSSD). Obtained as the square root of the mean of the sum of the squares of differences between

adjacent intervals. It can be expressed as:

$$\text{RMSSD}_i = \sqrt{\frac{1}{n-1}\sum_{j=1}^{n}(RR_{j+1} - RR_j)^2}$$

- Pairs of adjacent RR intervals differing by more than 50 ms (pNN50). Computed as the the number of pairs satisfying the condition in the analysis period divided by the by the total number of all RR intervals. It is denoted as:

$$\text{pNN50}_i = P(|RR_{j+1} - RR_j| > 50ms)$$

On the other hand, the following feature is computed in the time domain:

- LF/HF ratio (Low Frequency/High Frequency). It is computed as the ratio between the power in the low frequency LF_{HRV} and high frequency HF_{HRV} frequency bands. It is computed following the expression:

$$\text{LF/HF ratio}_i = \frac{\int_{0.04Hz}^{0.15Hz} f(\lambda)d\lambda}{\int_{0.15Hz}^{0.40Hz} f(\lambda)d\lambda}$$

where $f(\lambda)$ spectrum of the RR. This feature reflect the degree of sympathovagal balance with a higher ratio.

4.2 Electrodermal Activity

This feature extraction model employs our previously presented GSR feature extraction method available at [32]. This model divide the GSR signal into two components:

- **Skin Conductance Lelvel (SCL)**: This is related to the level of attention of the subject. If a person is concentrated and/or involved in a task, an increasing slope should be observed [29]. Otherwise, the SCL slope should either decrease or remain constant.

- **Skin Conductance Response (SCR)**: This is the indicator of sympathetic reactions (smaller than 10 seconds). This method allows to extract

their locations and durations. Furthermore, the sparsity of the resulting signal enhances its interpretability, whereas the post-processing stage allows to avoid false alarms.

The two components are computed for each $i-th$ interval obtaining the interval tonic component SCL_i and interval phasic component SCL_i. Given those components the following features can be computed:

- SCL slope.

$$m_i = \frac{SCL_{i+1} - SCL_i}{t_{i+1} - t_i}$$

  where SCL are the values of the slope for the i^{th} interval and t is the elapsed time (in this case are always 1 minute).

- SCR: peaks on an interval.

$$SCR_i = \sum(d_i > 0)$$

  where d is the driver obtained for each i^{th} interval (non-zero values are sympathetic reactions).

- AUC: area under the curve.

$$AUC_i = \sum d_i$$

  and finally values of the driver for each i^{th} interval are summed.

4.3 Speech Features

There are different ways to capture information contained in speech relevant to stress detection. In this paper, a simplified feature set is used, based on the InterSpeech 2009 Emotion [65] and Paralinguistic Challenge [66] features. The features are extracted with the openSMILE toolbox [20].

Five low-level descriptors of voice quality and pitch are computed: the estimated pitch, the voicing probability, the Jitter and the Shimmer. Finally the smoothed energy computed using the overlapping time frames is also included.

The pitch is computed as the envelope of the smoothed fundamental frequency contour obtained from the Cepstrum, the voicing probability is also obtained using the same estimated fundamental frequency. The Jitter and Shimmer are computed as the local (frame-to-frame) Jitter (pitch period length deviations) or Shimmer (amplitude deviations between pitch periods).

Finally in order to combine the speech features with the previously presented hearth rhythm and electrodermal activity features the arithmetic mean of the features contour is performed in a 60 second basics.

4.4 Cortisol Hormone

A cortisol test is performed by the participant right before and after of the experiment realization. The difference between obtained values provides a feature of the stress levels subjected by the participant during the experiment.

5

Stress Modeling

This chapter presents three different experiments developed in controlled situations where stress is incited to participants: 1) first experiment consists on 24 subjects playing neurocognitive games of increase difficulty mixed with relax periods, 2) second proof represents the motorization of individuals while are discussing a public talk, and finally 3) third case is a complete system created to extract features and classify stress level in a five-start raking scale.

These three implementations use some features discussed in previous section. Finally, some conclusions are reviewed to finish the implementation of this stress level improvement.

5.1 Physiological Feature Extraction In Neurocognitive Games

Neurocognitive games have recently been introduced in our daily life with the evolution of the smart-phones technology. Who has not played *tetris* while being in the subway? or have escaped a stressful day playing five minutes before going to sleep?. *Candy Crash* game, became the first ever game to be number 1 on iOS, Android and Facebook at the same time, and his slogan is 'play five minutes and remove your stress day'. But, do neurocognitive games really decrease our stress as their marketing says?. Which are the physiological features varying when we change our emotions?.

Emotional states provoke changes in physiological signals and could be acquired in non intrusive a way to obtain information about the mental state of the individual. The problem of recording these signals in real-time is that a supervised grand truth is needed to understand witch is the subjects state. Some authors have used neurocognitive games as ground truth to understand these human behaviors in a subjective way. P. Renaud and J. Blondin [63] used *Stroop color* test [39], a recognition test with in congruent questions to obtain the individual performance based on number of correct answers. They concluded that a solid relation exists but this test does not overcome an individual characterization. E. Jovanov et al. [40] used specific military personal training to know the performance of a soldier based on time reaction to achieve a mission and correlated it with physiological responses, obtaining a high correlation, but was not rating in a scale to measure it.

This experiment continues last authors works, performing a open schedule composed of neurocognitive games mixed with relaxing slots. The main objectives of this work are:

- To determine relevant physiological features which predict individual performance based on neurocognitive games.
- To classify between different states (relaxing or playing games) using the

most relevant features.

- To analyze the learning capabilities of a specific subject and show the differences between them.
- To compare between the subjective impression of each subject and the objective results of the physiological features extracted.

The overview of this experiment is divided into three main steps: first one is a feature extraction algorithm applied over the heart rhythm, Galvanic Skin Response (GSR) and hormone cortisol signals, obtaining relevant features. Secondly, the experiment design and a multi-classification stage used to determine the performance state. Finally, a individual learning criteria is applied to compare objective impressions versus the physiological results obtained.

5.1.1 Feature Selection Criteria

The duration of each experiment session is 106 minutes. Each feature is extracted using a period of one minute. The following eight stress features have been extracted from each record:

- Heart rhythm features:
  - SDNN,
  - RMSSD,
  - pNN50,
  - LF/HF ratio.
- Electrodermal activity:
  - slope of the signal,
  - area under the curve of the driver,
  - number of sparse driver activations.
- Hormone analysis:

– cortisol hormone (linear line regression between both points).

These features attempts to model the relationship between variables $x_{8\times106}$ by fitting a linear regression model to the observer data (scores obtained) $s_{1\times106}$. LASSO (Least Absolute Shrinkage and Selection Operator) [77] algorithm assigns different weights $w_{1\times8}$ to each of the input variables depending on the information provided by each one.

The selection of the most relevant features is calculated using the Mean Square Error (MSE) between normalized scores and the linear regression obtained. $MSE_j = \sum_{i=1}^{N}(s_j - w_i x_i)^2$, where i is the iteration of the $i-th$ feature leave out of the MSE, N is the number of features and i is the subject. The MSE is calculated using the method Leave-One-Out (LOO) so this work find the minimum MSE obtained to know which features contribute more.

5.1.2 Experiment Design

Classification between relaxing and playing games

The objective is to classify only using the most relevant features obtained in the previous subsection. The idea is to perform a binary classification achieving the maximum accuracy using a cross validation. Proposed classification methods were: binary decision trees, regularized linear and quadratic discriminant analysis, naive Bayes model with Gaussian, k-nearest neighbors classification, support vector machine and random forest. Secondly, the problem varies in a six multi-class support, where the same methods are used to classify between: relax time, speed, memory, attention, flexibility and problem solving tests.

Individual learning criteria

Subjects perform the experiment twice in order to know the capacity of learning the repetition of the same action one week later. A priori, scores should be higher the second time, since taking into account that the subject already knows the game and his ability to overcome himself. The technique to evaluate this is a

simple ratio, where: $learning\ ratio_i = \overline{Score\ 2nd\ time_i}/\overline{Score\ 1st\ time_i}$, where i is the category game or relax time and $\overline{x}$ denote the mean average.

Objective impression versus physiological subjectives results

The experiment consists of a set of neurocognitive tasks that measure five aspects: speed, memory, attention, flexibility and problem solving abilities. These tasks are developed by *Lumosity labs* and this experiment is composed by two games for each category (a total of ten games per experiment).

- Speed test: the intended primary skill targets are brain processing speed and participant reaction time.
- Memory tests: they are categorized under memory brain area, focusing on short-term memory.
- Attention tests: they are selective attention combined with multitasking.
- Flexibility tests: they are meant to determine mental skills to process multiple tasks simultaneously and the ability to switch between them.
- Problem solving tests: participants are expected to work through the details of a problem to reach a solution.

Finally, each participant fill a subjective questionnaire where they are asked what performance they think they have obtained from 1 to 5, where 1 was the worse and 5 the best performance. A multinomial logistic regression model was fit for each participant. The output was an integer number from 1 to 5, trying to compare these objective magnitudes with the subjective answers for each participant.

Relax time vs neurocognitive games

Experiment consists on an alternating set of games followed by a 150 seconds of participant relax. The hold time is 106 minutes. A representation of the schedule is shown in Fig. 5.1. It has been completed on a set of 24 participants with age

R	1	R	2	R	3	R	4	R	5	R	6	R	7	R	8	R	9	R	10	R

Figure 5.1: Experiment overview, where integer numbers represent the slot of each test and letter 'R', relax time. Note that number 1 and 6 belong speed tests, and so on: memory tests are 2 and 7, attention 3 and 8, flexibility 4 and 9, and problem solving tasks are 5 and 10.

interval from 20 to 59, mean age of the group was 29.16 years, with a standard deviation of 33.21. All subjects have undergone the experiment on a voluntary basis.

Data acquisition

Two non-intrusive sensors were used to measure physiological signals. The devices selected have been chosen from a wide range of available commercial sensors due to their reliability and small size.

Heart rhythm features acquired with *Microsoft Band 2* [55]: is a compact wristband that allows obtaining data about GSR, heart rate, Inter-Beat (RR) interval, device position and angular velocity.

Electrodermal activity recorded with *Q sensor* [1]. *Q sensor* allows the measurement of galvanic skin activity through two silver electrodes placed on the base of the sensor.

Individual Performance Every person has different responses in presence of difficult situations, but present similar behaviors performing the same activity. In this experiment, individual performance is simulated from different scores obtained on games with the given reaction time and to the correct answers of the tasks.

5.1.3 Results

#	(1) Relax vs games [Pe]			(2) Classification between games [Pe]							(3) Individual Learning averages [ratio]						(4) Comparative		
	Re	Ga	T	g1	g2	g3	g4	g5	Re	T	g1	g2	g3	g4	g5	T	Qu	Cl	MSE
1	30.77	6.25	12.26	75.00	62.50	87.50	43.75	50.00	53.85	61.32	1.22	1.31	1.45	1.34	1.88	1.44	4	1	9
2	23.08	7.50	11.32	50.00	50.00	62.50	81.25	50.00	57.69	58.49	0.45	0.32	0.80	0.70	1.10	0.67	2	3	1
3	23.08	7.50	11.32	50.00	81.25	56.25	37.50	75.00	42.31	55.66	1.20	1.34	1.34	1.56	2.23	1.53	4	4	0
4	30.77	10.00	15.09	43.75	18.75	31.25	62.50	43.75	96.15	53.77	1.60	1.22	1.76	1.22	4.20	2.00	5	5	0
5	30.77	6.25	12.26	93.75	75.00	56.25	68.75	62.50	50.00	66.04	2.02	0.43	0.65	0.54	5.30	1.79	4	3	1
6	26.92	7.50	12.26	50.00	50.00	50.00	75.00	43.75	57.69	54.72	1.30	0.45	0.43	0.98	1.11	0.85	3	3	0
7	38.46	8.75	16.04	75.00	75.00	93.75	43.75	56.25	42.31	62.26	2.51	1.22	1.23	1.03	1.00	1.40	3	4	1
8	19.23	7.50	10.38	50.00	93.75	31.25	50.00	81.25	76.92	65.09	2.01	1.56	1.92	1.09	1.09	4.2	5	4	1
9	38.46	11.25	17.92	75.00	31.25	50.00	81.25	62.50	42.31	55.66	1.29	1.40	1.91	1.34	1.20	1.43	3	3	0
T	29.06	8.06	13.31	62.50	59.72	57.64	60.42	58.33	57.69	59.22	1.51	1.03	1.28	1.09	2.47	1.48	3.67	3.33	1.44

Table 5.1: Statistical report for each subject(#): (1) probability of error of binary classification between relax and games. (2) probability of error of a multi-class classification. (3) ratio between first and second experiment. (4) Comparative between subjective appreciation and objective regression

An example of the signals recorded is shown in Fig. 5.2, as well as the scores obtained for half part of the experiment (53 minutes).

A Least Absolute Shrinkage and Selection Operator (LASSO) regression of the extracted features was fulfilled, using $\beta = 0.01$. Fig. 5.3 shows an example of how this regression adjust to the normalized scores, appearing rise during performed tasks and falls while rest periods.

Performing MSE function cost and the LOO method called before, the most relevant features that adjust better the regression was: (1) area under the curve of galvanic skin responses, and (2) LF/HF ratio and (3) Standard Deviation of NN Intervals (SDNN) of heart activity.

Classification results are shown in Table I, where appear the probability of error for each subject, between relax (Re) and games (Ga), and finally a the totals (T). This results were performed using the best accuracy option, a cross validation of Bayes naive with 64 neighbors and euclidean distance applied. Multi-classification was also calculated using a multi-modal logistic regression, optimizing the residual. In this case, as shown in Table I, it has higher errors differencing between games (g1, g2, g3, g4 and g5) and relax time (Re).

Learning averages are displayed in Table I, showing a statistical reporting for each subject. Finally, a rational quadratic regression was implemented to obtain an objective performance (Cl) and compare with the subjective (Qu) obtaining MSE between both.

5.1.4 Conclusions

The main goal of this experiment is to determine which physiological features are the most relevant to predict individual cognitive performance. In this open work, a statistical analysis of the predictive capacity of physiological signals is discussed.

LASSO regression algorithm combined with a cost function determines that the most relevant features were: area under the curve of electrodermal activity, LF/HF ratio and SDNN of cardiac activity. These three features are the most used in stress detection, and this work demonstrates that these can also be useful.

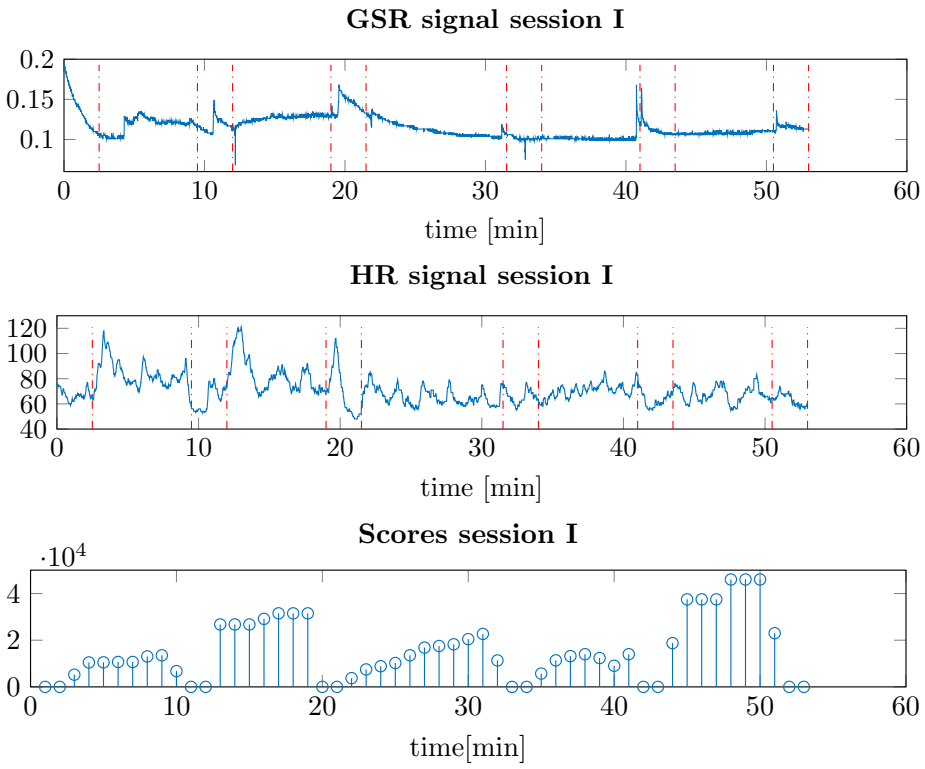


Figure 5.2: Example of data collected for half session of the experiment (53 minutes). Vertical red lines represent the differentiation between relax times and test. First subfigure (upper) shows GSR, Second (medium) the Heart Rate recorder and last figure (lower) the scores obtained.

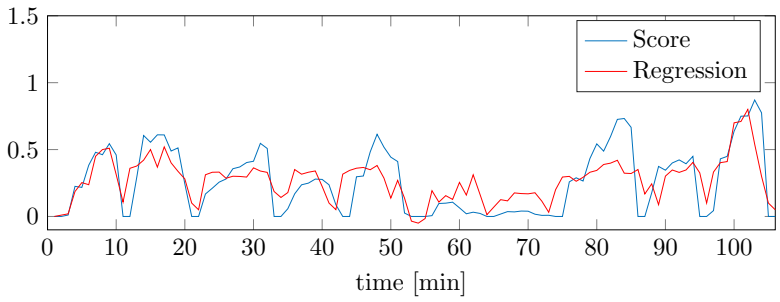


Figure 5.3: Example of a linear regression. Line blue represent the scores obtained for the participant (zero in relax times) and in color red the regression obtained using physiological signals.

In this experiment, the binary classifier distinguish between play games and relax slots achieving a 13.31% probability of error. A multi-modal classifier performed more poorly, achieving only 59,22% mean square error where it is not possible to correlate the type of game with physiological features.

Moreover, results of nine subjects has been shown in Table I where each one has different capabilities. Besides, on average, they have achieved a 1.48 learning ratio so it can be said that each participant learned while playing, and achieved better results once the games were already familiar to them.

Depending on the game, the learning time was different of the game, capability of learning change, for example, problem solving task has a ratio of 2.47 on average instead of 1.03 in memory test.

Finally, a regression fit physiological features in a subjective way, and correlate with the subjective appreciation of each person, reaching only a 1.44 MSE on average, as shown in Table I. This mean that the appreciation of each volunteers seem to be what his physiological reactions mean.

Taking this work as a reference, some future lines of research can be followed to improve the development of a complete system of stress detection to know our capabilities in real time.

5.2 Stress States Classification In Public Talks

Do you remember the last time you give a public talk? How did you feel when you start? Did you feel nervous in order to perform an impressive talk?. These affective symptoms that appear can suppose the appearance of stress [28].

Emotional symptoms provoke changes in different physiological signals that can be monitored to detect stress. D. Wu et al. [83] captured physiological responses: GSR, respiration, electroencephalogram and electrocardiogram for identification and classification into several stress levels. Another case was Zhai and Barreto [86], who acquired different affective features: GSR, blood volume pressure and pupil diameter to differentiate states (relax or mental stress) in computers user. Speech also plays a role in detecting stress as can be depicted form [16] where stress

is estimated using pitch and speech energy features and [58] where Log Frequency Power Coefficients (LFPC) are used to detect stress and emotion.

The main scope of this experiment is to extract physiological features of monitoring subjects discussing public talks and classify their stress levels. A novel dataset composed of 17 subjects is presented where speech, electrodermal activity and heart rhythm were recorder using non-intrusive sensors. A proposed classifier determine the status of 8 volunteers of the dataset: [A] pre-presentation period, [B] talk, [C] questions. The differentiation between theses three classes are imposed to replicate the protocol used in Trier Social Stress Test (TSST) test. Another authors also used this three intervals to: show stress activations induced cortisol levels [82], stress detection from speech using Galvanic Skin Responses [42], or acute stress and its relation with chronic trapezius myalagia [75].

Each feature is extracted using a period of one minute. The following twelve stress features have been extracted from each record:

- Heart rhythm features:
  - SDNN,
  - RMSSD,
  - pNN50,
  - LF/HF ratio.
- Electrodermal activity:
  - slope of the signal,
  - area under the curve of the driver,
  - number of sparse driver activations.
- Speech:
  - estimated pitch,
  - voicing probability,
  - jitter,

- shimmer,
- smoothed energy.

The feature vector can be denoted as $x_{12\times D}$, where D is the duration of the speech in minutes and 12 are the number of features. Several multi-class supervised classification methods were tested: k-nn neighborhoods, multi-class logistic regression and decision trees. Also non-supervised methods were validated as the expectation-maximization algorithm (EM) and Gaussian Mixture Models with Dirichlet Process (GMM-DP).

The models are trained using 9 talks as training set and tested in the rest. Note that monitoring each talks needs the acceptance of each individual, storage all the data and process it which involves difficult to make bigger the database.

Finally, the algorithm proposed was an multi-class logistic classifier parametrized by a weight matrix and a bias vector (W,b). Classification is done by projecting data points onto a set of hyper-planes, the distance to which is used to determine a class membership probability. Mathematically this can be expressed as:

$$P(Y = i|x, W, b) = \frac{e^{W_i x + b_i}}{\sum_1 e^{W_j x + b_j}}$$

, corresponding to each class (y_i) logistic classifier is characterized by a set of parameters (W_i,b_i).

5.2.1 Experimental Set-up

The experiment has been completed on a set of 17 participants with age interval from 20 to 59, mean age of the group was 29.2 years, with a standard deviation of 33.21. All subjects have undergone the experiment on a voluntary basis.

Three non-intrusive sensors were used to measure physiological signals. The devices selected have been chosen from a wide range of available commercial sensors due to their reliability and small size. The different signal are acquired using the following devices and configurations:

- **Heart rhythm signal** is acquired with *Microsoft Band 2* [55], it is a compact wristband that allows obtaining data about galvanic skin responses

#	Classification between states				
	Duration (min)	A	B	C	Average
1	43	00.00	16.67	8.33	13.95
2	40	00.00	30.00	18.18	20.00
3	57	00.00	8.51	20.00	1.75
4	66	20.00	17.02	11.11	16.67
5	58	00.00	12.90	57.89	25.86
6	66	00.00	14.29	22.22	15.15
7	63	50.00	14.89	14.29	15.87
8	63	00.00	12.50	12.50	11.11
Average	57	15.85	8.06	20.57	15.05

Table 5.2: Statistical report for each subject(#): probability of error of a multi-class classification. Each column represent: (1) speaker identifier, (2) Duration of the talk, (3) Probability of error before talk [A], (4) Probability of error during talk [B], (5) Probability of error in question time [C] and (6) Total average.

(GSR), heart rate, RR interval, device position and angular velocity.

- **Electrodermal activity** is recorded with a *Q sensor* [1]. It allows the measurement of galvanic skin activity through two silver electrodes placed on the base of the sensor.

- **Speech** is obtained using a Zoom H1 handheld recorder with a Rode lavalier microphone. Non lossy compressed files (wav), 24-bit quantification and 44 kHz sampling rate are used.

The captured signals and the computed features can be download form [31].

5.2.2 Results

Multi-class logistic regression algorithm achieves the better accuracy of the models raised before. Table 1 display the probability of error achieved for classification obtained. It is displayed the duration of the talks and the probability of error for each time division: probability of error pre-presentation period [A], probability of

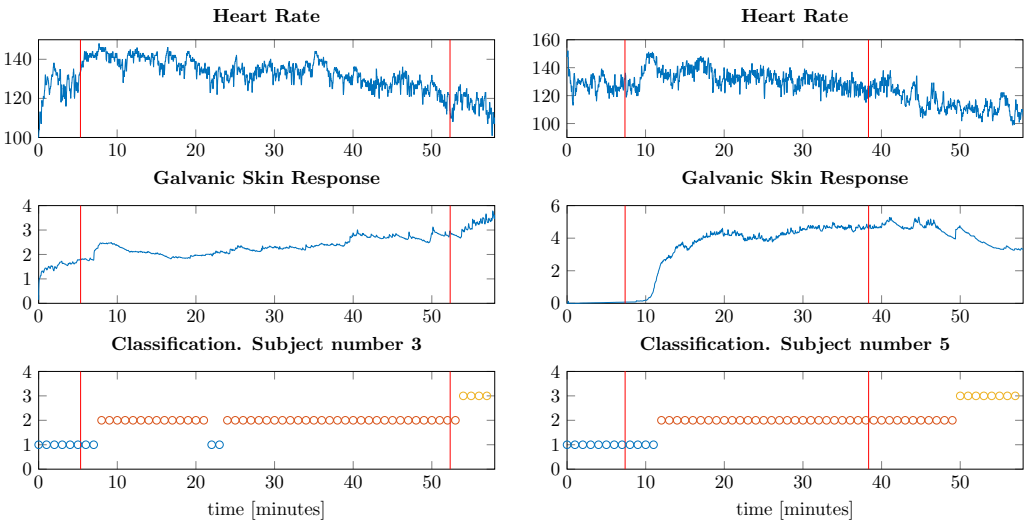


Figure 5.4: Examples of data classified of the experiment. Vertical red lines represent the differentiation between states: before talk [state A], talk [state B] and questions [state C]. Left figure corresponds to number 3 subject and right to number 5. First sub-figure (upper) shows Heart Rate, Second (medium) the Galvanic Skin Response and last figure (lower) the classification obtained in three classes: blue, red and yellow.

error during talk [B], probability of error in question time [C]. Besides an average column is included for further conclusions. Fig. 2 shows two examples of the heart rate, galvanic skin response and the classification obtained for the subjects number 4 and number five displayed in Table 5.2.

5.2.3 Conclusions

This work proposes a new framework to classify stress states using only physiological signals. 17 volunteers were monitored while discussing a public talk. 9 of them were used as training set and the rest were used as test set.

Model proposed was selected among many others used because obtain the best accuracy no over-fitting. Table 1 represents the multi-class classification obtained using a multi-class regression algorithm. Results show a 15.05% probability of error in average that confirms stress features were well chosen and model achieve the expectations.

Two examples of the recorded signals and their stress classification were dis-

played in Fig. 2 to understand the model proposed. The classifier also give the more relevant features (W): the slope of the GSR signal, the LF/HR ratio of heart rhythm and the estimated speech pitch.

5.3 Real-time Stress Classification

Daily work supposes the first factor of stress as it was mentioned in this thesis introduction [64]. In fact, depends on the type of work can be more dangerous, need a deal of concentration, questions of the boss could imply an extra difficulty, anxiety situations, etc. Therefore a wrong decision can bring serious consequences. Besides, also depends on the individual performance for each person: how reacts the worker in presence of pressure or indecision situations, and if s/he is trained and qualified to this end.

The aim of this experiment is to monitor a Unmanned Aerial Vehicle (UAV) operator in his/her work environment and determine in real-time his/her stress level. These kind of workers are trained in pressure situations, so first the system should characterized each individual and them, test it in a real work environment situation. The operator use to be sitting but also use to move around the UAV controller so the system acquisition should acquire the signals in a non-intrusive way.

The hold system analyzes physiological signals captured and the output is a level of stress from one to five, where one seem the most relaxed and five the most stressed. Another additional outputs could be: no signal (NS) and died (D).

This work is subdivided in three parts: individual stress characterization for each operator, a complete real-time system for stress classification and some results and conclusions obtained.

5.3.1 Individual stress characterization

Two operators of UAV (named Sofia and Oscar) were asked in voluntary basis to fulfill the neurocognitive games presented in Section 5.1.2. They completed the experiment twice, with a month of difference between them.

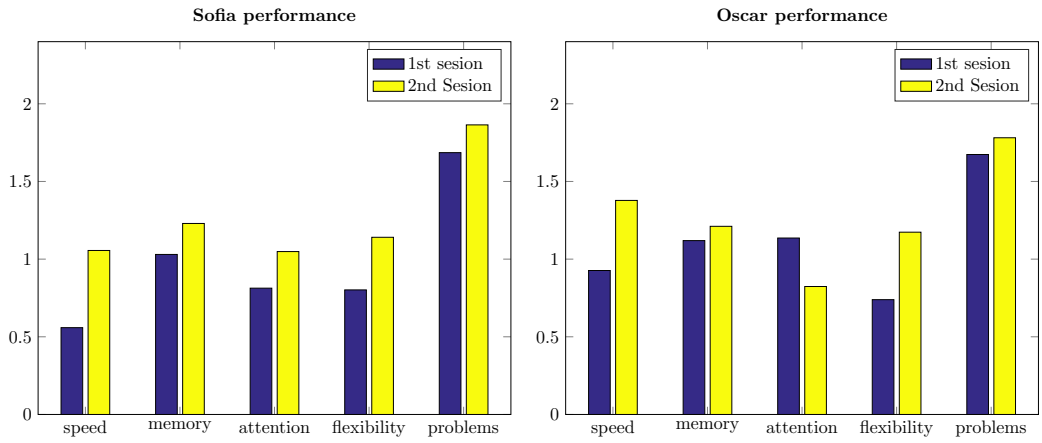


Figure 5.5: Comparative between the scored obtained in the first and second sessions.

The games could be clustered in five abilities: speed, memory, attention, flexibility and problem solving tests. Figure 5.5 presents the difference between first and second sessions in average for each participant. As are scores normalized, the capacity of learning can be extracted for each participant and also know if they are over/under the global mean (normalized as 1).

Second question is to characterize the Yerkes & Dodson curve named in chapter 1 and can be found in Figure 1.1. Each test is normalized and of increased complexity. Also taking the average for all the tests, Figure 5.6 represents the mean average scores normalized by difficulty. These curves shows the level of difficulty from 1 to 9 and the y-axis represents the level of performance. Both representations show a similar inverted-U as the YerKes & Dodson curve.

5.3.2 Complete system

A complete real-time system was developed for stress level assessment. The scheme of this system, that can be followed in Figure 5.7, is composed of:

- A non-intrusive sensor, Microsoft Band 2, that monitor heart rhythm and electrodermal activity signals, and send the values and their correspondent timestamps to a threat installed in a computer.

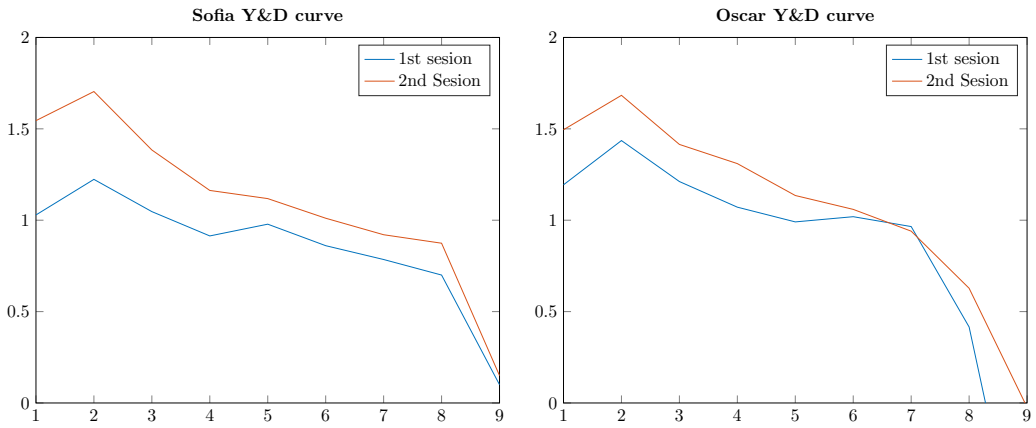


Figure 5.6: Simulated Yerkes & Dodson curves for each participant and session.

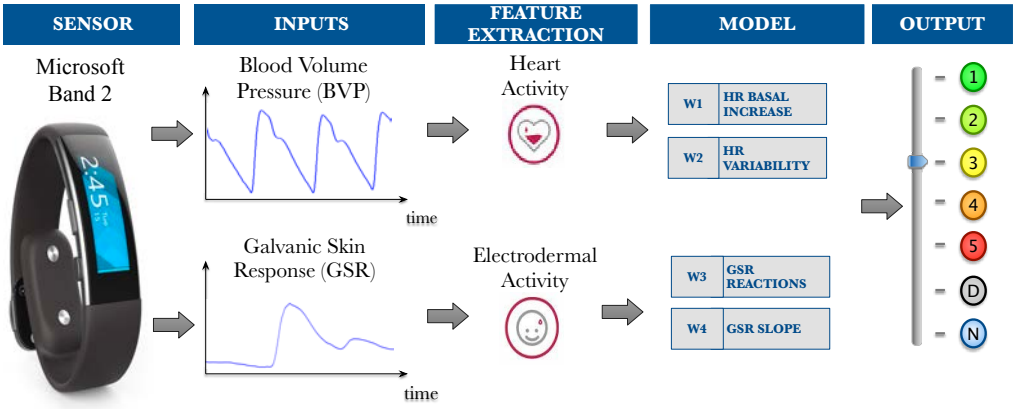


Figure 5.7: Graphical scheme of the real-time acquisition, signal processing and stress level classification.

- A threat developed in *Visual Studio* is waiting each second for a package send from the sensor via *bluetooth*. Once it receive the values in form of package, the program save them in a buffer.

- Another threat in parallel send the package from *Visual Studio* to *Matlab*.

- *Matlab* receives each second a package and process it in frames of 60 seconds.

- The final output is a stress level assessment closed in a five-star rating scale.

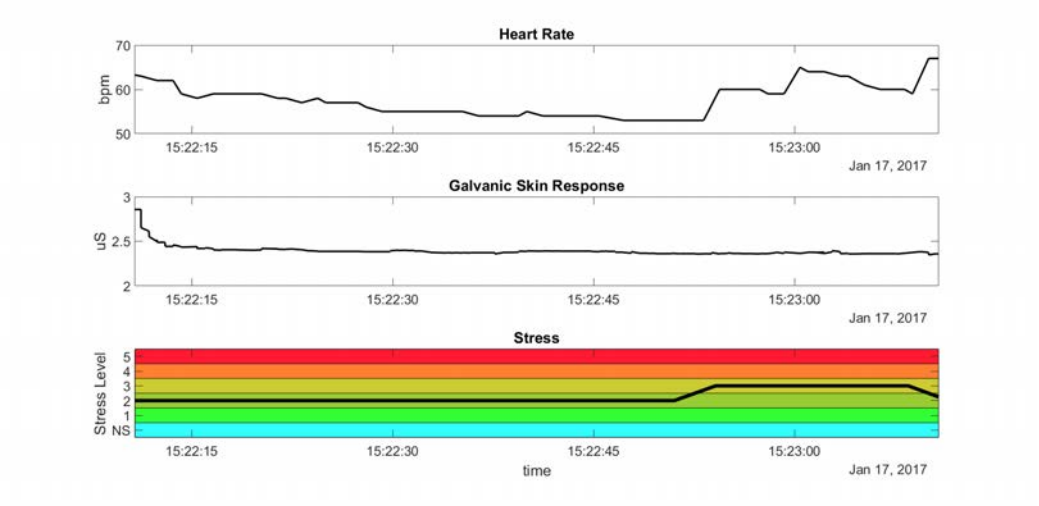


Figure 5.8: Example of real-time demonstrator report of 60 seconds.

5.3.3 Results

Two different reports are developed. The first one is real-time report of 60 seconds that display the Heart Rate and GSR signals and the correspondent level of stress. An example is displayed in Figure 5.8.

The second report is an overview for an session finished. Figure 5.9 shows an example of an record of 15 minutes while an operator was working.

In general for every work session worked use to be at level 3 that implies that a normal performance. In some critical slot time the level of stress increases because of lot of warning or a critical problem. In the other hand, the levels could decrease from 3 if the UAV is totally controlled or stooped.

5.3.4 Conclusions

In this experiment two objectives have been achieved. First one is to characterize an individual because the reaction of each person are different even performing the same activity. This experiment presents a controlled test where the capacities for each volunteer in five different abilities: speed, memory, attention, flexibility and problem solving tests. Besides, Figure 5.5 is normalized to 1 where respect to a global average mean. An individual Yerkes & Dodson curve is also defined in

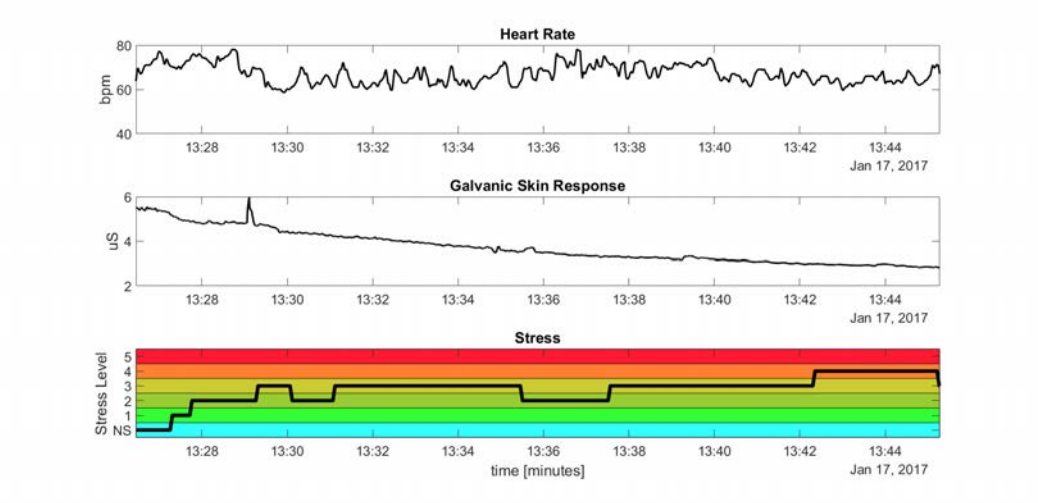


Figure 5.9: Example of a report of an experiment of 15 minutes.

term of the level difficulty that is shown in Figure 5.6.

On the other hand, a real-time classifier was implemented based on Heart Rate and Electrodermal activity. A *Microsoft Band 2* captured the signals and send it to a computer that analyze and decide the level of stress. This system is a complete real-time non-intrusive level assessment that cloud be useful to manage the workload for each operator. Basides, this method can be extrapolated to others critical workers, such as, comercial flight pilots, police, etc.

6 Conclusions

6.1 Summary

The main objective of this thesis is the development of novel stress assessment employing the raw signals provided by wearable sensors. In this chapter, contributions of this thesis are summarized, and some future research lines are described.

- In Chapter 2, the state of the art of stress assessment is presented. A stress models history, an introduction of the Autonomous Nervous System (ANS) and possible objectives measurements are discussed to overcome contributions of this thesis. Besides, the sensors used during this thesis are included and a novel of possible physiological signals for stress reactions detection.
- In Chapter 3, a new feature extraction method for Galvanic Skin Response (GSR) signals is presented. Past methods are not interpretable to apply

statistical feature extraction and this improvement employs an sparse model that clarify the output obtained. Furthermore, the method is faster than others and can be implemented in a wearable sensor due to its low need of signal processing. The contributions of this chapter were presented in [35].

- In Chapter 4, most usual signal processing methods are presented to extract stress features. It includes the cited algorithm [32] of GSR responses, and some common methods for heart rhythm, speech and cortisol hormone analysis. All of them compose a complete feature extraction model that can be applied depending on the work or experiment requirements.

- Finally, Chapter 5 defines three experiments implemented to validate the feature extraction model proposed.

  - The first one captured signals and extract features in a controlled environment while the subject play neurocognitive games. We can conclude which were the most relevant features using a linear regression: area under the curve of electrodermal activity, LF/HF ratio and Standard Deviation of NN Intervals (SDNN) of cardiac activity.

  - Second experiment is a controlled environment of public talks. It shows a percentage of 15.05% probability of error in average while differentiate between talk, before and after talks, using the feature extraction model proposed in Chapter 4.

  - The last experiment shows a complete system that individualize stress reaction evaluating the model proposed. The system captured in real-time physiological signals and classify in a five levels rating scale.

These three implementations consolidate the objectives proposed for this thesis focusing on stress level assessment.

6.2 Future Lines

This work also suggests several paths for further research in the stress modeling system with wearable sensors. We provide below a list with what we consider are some potential research lines.

Real-time stress bracelet. This thesis has overcome a feature extraction model for stress assesment and could be implemented in a wearable sensor. A proposed application of this method should be a linear classifier that notify the level of stress in real-time. The bracelet should include a display where shows a number normalized in a scale as a commercial wearable.

Workload assessment. In a work environment, it should be useful a non-intrusive system to know the level of stress to increment or decrement the workload in time. This technique can be implemented in critical works such as: operator of Unmanned Aerial Vehicle (UAV), police, flight pilots, firefighters, etc, depending on the need.

References

[1] Affectiva. Qsensor. `http://qsensor-support.affectiva.com/`.

[2] Jordi Aguiló, Pau Ferrer-Salvans, Antonio García-Rozo, Antonio Armario, Ángel Corbí, Francisco J Cambra, Raquel Bailón, Ana González-Marcos, Gerardo Caja, Sira Aguiló, et al. Project ES3: Attempting to quantify and measure the level of stress. *Revista de neurologia*, 61(9):405–415, 2015.

[3] D. M. Alexander, C. Trengove, P. Johnston, T. Cooper, J. P. August, and E. Gordon. Separating individual skin conductance responses in a short interstimulus-interval paradigm. *Journal of Neuroscience Methods*, 146(1):116–123, 2005.

[4] John Allen. Photoplethysmography and its application in clinical physiological measurement. *Physiological measurement*, 28(3):R1, 2007.

[5] Guillermo N Armaiz-Pena, Susan K Lutgendorf, Steve W Cole, and Anil K Sood. Neuroendocrine modulation of cancer progression. *Brain, behavior, and immunity*, 23(1):10–15, 2009.

[6] Armando Barreto and Jing Zhai. Physiologic Instrumentation for Real-time Monitoring of Affective State of Computer Users. *WSEAS Transactions on Circuits and Systems*, 3:496–501, 2003.

[7] Armando Barreto, Jing Zhai, and Malek Adjouadi. Non-intrusive physiological monitoring for automated stress detection in human-computer interaction. *Human–Computer Interaction*, pages 29–38, 2007.

[8] Mathias Benedek and Christian Kaernbach. A continuous measure of phasic electrodermal activity. *Journal of Neuroscience Methods*, 190(1):80–91, 2010.

[9] Mathias Benedek and Christian Kaernbach. Decomposition of skin conductance data by means of nonnegative deconvolution. *Psychophysiology*, 47(4):647–658, 2010.

[10] Wolfram Boucsein. *Electrodermal activity*. Springer Science & Business Media, 2012.

[11] M M Bradley and P J Lang. Handbook of psychophysiology. *Handbook of psychophysiology*, 2000.

[12] Margaret M Bradley and Peter J Lang. Handbook of psychophysiology, 2007.

[13] Phillip J Brantley, Craig D Waggoner, Glenn N Jones, and Neil B Rappaport. A daily stress inventory: Development, reliability, and validity. *Journal of behavioral medicine*, 10(1):61–73, 1987.

[14] John T Cacioppo, Gary G Berntson, Jeff T Larsen, Kirsten M Poehlmann, Tiffany A Ito, et al. The psychophysiology of emotion. *Handbook of emotions*, 2:173–191, 2000.

[15] A. John Camm and Marek Malik. Heart rate variability. *European Heart Journal (1996)*, pages 354–381, 1996.

[16] L. Czap and J. M. Pintér. Intensity feature for speech stress detection. In *Proceedings of the 2015 16th International Carpathian Control Conference (ICCC)*, pages 91–94, May 2015.

[17] Bradley Efron, Trevor Hastie, Iain Johnstone, and Robert Tibshirani. Least angle regression. *The Annals of statistics - Institute of Mathematical Statistics*, 32(2):407–499, 2004.

[18] Michael Elad. Sparse and redundant representations: From theory to applications in signal and image processing, 2010.

[19] Empatica. Empatica E4 wristband. `https://www.empatica.com/e4-wristband`.

[20] Florian Eyben, Felix Weninger, Florian Gross, and Björn Schuller. Recent developments in opensmile, the munich open-source multimedia feature extractor. In *Proceedings of the 21st ACM International Conference on Multimedia*, MM '13, pages 835–838, New York, NY, USA, 2013. ACM.

[21] Christos A Frantzidis, Evdokimos Konstantinidis, Costas Pappas, and Panagiotis D Bamidis. An automated system for processing electrodermal activity. *Studies in health technology and informatics*, 150:787, 2008.

[22] Hugues Garnier and Liuping Wang. *Identification of Continuous-time Models from Sampled Data*. Springer Science & Business Media, 2008.

[23] Alberto Greco, Gaetano Valenza, Antonio Lanata, Enzo Pasquale Scilingo, and Luca Citi. cvxeda: A convex optimization approach to electrodermal activity processing. *IEEE Transactions on Biomedical Engineering*, 63(4):797–804, 2016.

[24] John E Hall. *Guyton and Hall textbook of medical physiology*. Elsevier Health Sciences, 2010.

[25] Jennifer A Healey. Wearable and automotive systems for affect recognition from physiology. *(Doctoral dissertation, Massachusetts Institute of Technology)*, 2000.

[26] Jennifer A Healey and Rosalind W. Picard. Detecting stress during real-world driving tasks using physiological sensors. *Intelligent Transportation Systems, IEEE Transactions on*, 6(2):156–166, 2005.

[27] Jennifer A Healey and Rosalind W Picard. Detecting stress during real-world driving tasks using physiological sensors. *IEEE Transactions on intelligent transportation systems*, 6(2):156–166, 2005.

[28] Javier Hernandez, Pablo Paredes, Asta Roseway, and Mary Czerwinski. Under Pressure: Sensing Stress of Computer Users. *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14*, pages 51–60, 2014.

[29] Javier Hernandez, Ivan Riobo, Agata Rozga, Gregory D Abowd, and Rosalind W Picard. Using electrodermal activity to recognize ease of engagement in children during social interactions. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 307–317. ACM, 2014.

[30] Alberto Hernando, Jesús Lázaro, Eduardo Gil, Adriana Arza, Jorge Mario Garzón, Raúl López-Antón, Concepción de la Cámara, Pablo Laguna, Jordi Aguiló, and Raquel Bailón. Inclusion of respiratory frequency information in heart rate variability analysis for stress assessment. *IEEE journal of biomedical and health informatics*, 20(4):1016–1025, 2016.

[31] Hernando-Gallego and F. de la Calle Silos. Public talks stress dataset. `http://www.tsc.uc3m.es/~fsilos/stress_dataset.zip`.

[32] F. Hernando-Gallego. Sparse eda. `https://github.com/fhernandogallego/sparsEDA`.

[33] F. Hernando-Gallego and A. Artés-Rodríguez. Individual performance calibration using physiological stress signals. *Workshop on IEEE Conference on Body Sensor Networks (2015)*, July 2015.

[34] F. Hernando-Gallego and A. Artés-Rodríguez. Physiological Feature Extraction in Neurocognitive Games. *IEEE Conference on Body Sensor Networks (2018)*, Under revision 2017.

[35] F. Hernando-Gallego, D. Luengo, and A. Artés-Rodríguez. Feature Extraction of Galvanic Skin Responses by Non-Negative Sparse Deconvolution. *IEEE Journal of Boimedical and Health Informatics*, pages 1–1, 2017.

[36] F. Hernando-Gallego, F. Silos, and A. Artés-Rodríguez. Stress States Classification Using Physiological Signals During Public Talks. *Acoustics, Speech, and Signal Processing (ICASSP) (2018)*, Under revision 2017.

[37] Eva Hudlicka. To feel or not to feel: The role of affect in human–computer interaction. *International journal of human-computer studies*, 59(1):1–32, 2003.

[38] Swayambhoo Jain, Urvashi Oswal, Kevin Shuai Xu, Brian Eriksson, and Jarvis Haupt. A compressed sensing based decomposition of electrodermal activity signals. *IEEE Transactions on Biomedical Engineering*, 64(9):2142–2151, 2017.

[39] Arthur R Jensen and William D Rohwer. The Stroop color-word test: A review. *Acta psychologica*, 25:36–93, 1966.

[40] Emil Jovanov, Amanda O. Lords, Dejan Raskovic, Paul G. Cox, Reza Adhami, and Frank Andrasik. Stress monitoring using a distributed wireless intelligent sensor system. *Engineering in Medicine and Biology Magazine*, 22(3)(June):49–55, 2003.

[41] Clemens Kirschbaum, K-M Pirke, and Dirk H. Hellhammer. The 'Trier Social Stress Test'–a tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology*, 28:76–81, 1993.

[42] Hindra Kurniawan, Alexandr V. Maslov, and Mykola Pechenizkiy. Stress detection from speech and Galvanic Skin Response signals. *Proceedings of the 26th IEEE International Symposium on Computer-Based Medical Systems*, pages 209–214, jun 2013.

[43] Sungjun Kwon, Jeongsu Lee, Gih Sung Chung, and Kwang Suk Park. Validation of heart rate extraction through an iphone accelerometer. In *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, pages 5260–5263. IEEE, 2011.

[44] Ledalab, University of Graz (Austria). GSR signal example. `http://ledalab.de/download/ivn07_16_matlab.mat`.

[45] Chong L. Lim, Chris Rennie, Robert J. Barry, Homayoun Bahramali, Ilario Lazzaro, Barry Manor, and Evian Gordon. Decomposing skin conductance into tonic and phasic components. *International Journal of Psychophysiology*, 25(2):97–109, feb 1997.

[46] David Luengo, Sandra Monzón, Tom Trigano, Javier Vía, and Antonio Artés-Rodríguez. Blind analysis of atrial fibrillation electrograms: a sparsity-aware formulation. *Integrated Computer-Aided Engineering*, 22(1):71–85, 2015.

[47] David Luengo, Javier Vía, Sandra Monzón, Tom Trigano, and Antonio Artés-Rodríguez. Cross-products LASSO. In *2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 6118–6122, 2013.

[48] V G Macefield, B G Wallin, and a B Vallbo. The discharge behaviour of single vasoconstrictor motoneurones in human muscle nerves. *The Journal of physiology*, 481 Pt 3:799–809, 1994.

[49] Vaughan G Macefield and B Gunnar Wallin. The discharge behaviour of single sympathetic neurones supplying human sweat glands. *Journal of the autonomic nervous system*, 61(3):277–286, 1996.

[50] Gloria Mark, Shamsi Iqbal, Mary Czerwinski, and Paul Johns. Capturing the mood: facebook and face-to-face encounters in the workplace. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*, pages 1082–1094. ACM, 2014.

[51] Andrea H. Marques, Marni N. Silverman, and Esther M. Sternberg. Evaluation of stress systems by applying noninvasive methodologies: Measurements of neuroimmune biomarkers in the sweat, heart rate variability and salivary cortisol. *NeuroImmunoModulation*, 17(3):205–208, 2010.

[52] Gerald Matthews, D Roy Davies, Rob B Stammers, and Steve J Westerman. *Human performance: Cognition, stress and individual differences*. Psychology Press, Philadelphia, 2000.

[53] Bruce S McEwen and Teresa Seeman. Stress and affect: Applicability of the concepts of allostasis and allostatic load. 2003.

[54] Medicom MTD Ltd. Medical equipment for neurophysiology, polysomnography, biofeedback and research. `http://www.medicom-mtd.com/en/`.

[55] Microsoft. Microsoft band 2. `https://www.microsoft.com/microsoft-band/`.

[56] Sandra Monzón, Tom Trigano, David Luengo, and Antonio Artés-Rodríguez. Sparse spectral analysis of atrial fibrillation electrograms. In *2012 IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*, pages 1–6, 2012.

[57] Paul Jarle Mork and Rolf H Westgaard. The influence of body posture, arm movement, and work stress on trapezius activity during computer work. *European journal of applied physiology*, 101(4):445–56, nov 2007.

[58] T. L. Nwe, S. W. Foo, and L. C. De Silva. Detection of stress and emotion in speech using traditional and fft based log energy features. In *Fourth International Conference on Information, Communications and Signal Processing, 2003 and the Fourth Pacific Rim Conference on Multimedia. Proceedings of the 2003 Joint*, volume 3, pages 1619–1623 vol.3, Dec 2003.

[59] Juha Pärkkä, Juho Merilahti, Elina M Mattila, Esko Malm, Kari Antila, Martti T Tuomisto, Ari Viljam Saarinen, Mark van Gils, and Ilkka Korhonen. Relationship of psychological and physiological variables in long-term self-monitored data during work ability rehabilitation program. *IEEE Transactions on Information Technology in Biomedicine*, 13(2):141–151, 2009.

[60] DH Phan, S Bonnet, R Guillemaud, E Castelli, and NY Pham Thi. Estimation of respiratory waveform and heart rate using an accelerometer. In *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, pages 4916–4919. IEEE, 2008.

[61] Rosalind W Picard and Roalind Picard. *Affective computing*, volume 252. MIT press Cambridge, 1997.

[62] Miguel RAMOS, C Rovira, L Umfuhrer, and E Urbina. Sistema nervioso autónomo. revisión. *Revista de Posgrado de la Cátedra VIa Medicina*, 1(101), 2001.

[63] Patrice Renaud and JP Blondin. The stress of Stroop performance: physiological and emotional responses Renaud, P., & Blondin, J. (1997). The stress of Stroop performance: physiological and emotional responses to color-word interference, task pacing, and pacing speed. International Jour. *International Journal of Psychophysiology*, pages 87–97, 1997.

[64] Paul J Rosch. The quandary of job stress compensation. *Health and Stress*, 3(1):1–4, 2001.

[65] Björn Schuller, Stefan Steidl, and Anton Batliner. A.: The interspeech 2009 emotion challenge. In *In ISCA, ed.: Proceedings of Interspeech*, pages 312–315, 2009.

[66] Björn Schuller, Stefan Steidl, Anton Batliner, Felix Burkhardt, Laurence Devillers, Christian Müller, and Shrikanth Narayanan. The interspeech 2010 paralinguistic challenge. In *In Proc. Interspeech*, 2010.

[67] Cornelia Setz, Bert Arnrich, Johannes Schumm, Roberto La Marca, Gerhard Tröster, and Ulrike Ehlert. Discriminating stress from cognitive load using a wearable eda device. *IEEE Transactions on information technology in biomedicine*, 14(2):410–417, 2010.

[68] Yuan Shi, Minh Hoai Nguyen, Patrick Blitz, Brian French, Scott Fisk, Fernando De la Torre, Asim Smailagic, Daniel P Siewiorek, Mustafa al'Absi, Emre Ertin, et al. Personalized stress detection from physiological measurements. In *International symposium on quality of life technology*, pages 28–29, 2010.

[69] Shimmer. *ExG User Guide for ECG*. www.shimmersensing.com, Realtime Technologies Ltd, 2014.

[70] Shimmer. *ExG User Guide for EMG*. www.shimmersensing.com, Realtime Technologies Ltd, 2014.

[71] Shimmer. *GSR + Expansion Board User Guide*. www.shimmersensing.com, Realtime Technologies Ltd, 2014.

[72] Shimmer. *Optical Pulse Sensing Probe User Guide*. www.shimmersensing.com, Realtime Technologies Ltd, 2014.

[73] Shimmer Sensing. Shimmer 3 GSR unit. `http://www.shimmersensing.com/products/shimmer3-wireless-gsr-sensor`.

[74] Mandeep Singh and AB Queyam. Stress Detection in Automobile Drivers using Physiological Parameters: A Review. *International Journal of Electronics Engineering*, 5(2):1–5, 2013.

[75] Anna Sjörs, Britt Larsson, Björn Karlson, Kai Osterberg, Joakim Dahlman, and Björn Gerdle. Salivary cortisol response to acute stress and its relation to psychological factors in women with chronic trapezius myalgia. *Psychoneuroendocrinology*, 35(5):674–85, jun 2010.

[76] Sara Taylor, Natasha Jaques, Weixuan Chen, Szymon Fedor, Akane Sano, and Rosalind Picard. Automatic Identification of Artifacts in Electrodermal Activity Data. *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE*, pages 1934–1937, 2015.

[77] Robert Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 267–288, 1996.

[78] Chadi Touma, Rupert Palme, and Norbert Sachser. Analyzing corticosterone metabolites in fecal samples of mice: a noninvasive technique to monitor stress hormones. *Hormones and Behavior*, 45(1):10–22, jan 2004.

[79] B Gunnar Wallin. Sympathetic nerve activity underlying electrodermal and cardiovascular reactions in man. *Psychophysiology*, 18(4):470–476, 1981.

[80] Quan Wang. COSBOS: COlor-Sensor-Based Occupancy Sensing. `https://se.mathworks.com/matlabcentral/fileexchange/48428-cosbos--color-sensor-based-occupancy-sensing/content/COSBOS_v1.0/LTM_Recovery/Lib/SparseLab2.1-Core/Solvers/SolveLasso.m`, 12 Nov. 2014.

[81] Elliot D Weitzman, David Fukushima, Christopher Nogeire, Howard Roffwarg, T F Gallagher, and Leon Hellman. Twenty-four hour pattern of the episodic secretion of cortisol in normal subjects. *The Journal of Clinical Endocrinology \& Metabolism*, 33(1):14–22, 1971.

[82] Oliver T. Wolf, Nicole C. Schommer, Dirk H. Hellhammer, Bruce S. McEwen, and Clemens Kirschbaum. The relationship between stress induced cortisol levels and memory differs between men and women. *Psychoneuroendocrinology*, 26(7):711–720, oct 2001.

[83] Dongrui Wu, Christopher G. Courtney, Brent J. Lance, Shrikanth S. Narayanan, Michael E. Dawson, Kelvin S. Oie, and Thomas D. Parsons. Optimal Arousal Identification and Classification for Affective Computing Using Physiological Signals: Virtual Reality Stroop Task. *IEEE Transactions on Affective Computing*, 1(2):109–118, jul 2010.

[84] Robert M Yerkes and John D Dodson. The relation of strength of stimulus to

rapidity of habit-formation. *Journal of comparative neurology and psychology*, 18(5):459–482, 1908.

[85] Jing Zhai and Armando Barreto. Stress detection in computer users based on digital signal processing of noninvasive physiological variables. In *Engineering in Medicine and Biology Society, 2006. EMBS'06. 28th Annual International Conference of the IEEE*, pages 1355–1358. IEEE, 2006.

[86] Jing Zhai, Armando Barreto, Craig Chin, and Chao Li. Realization of stress detection using psychophysiological signals for improvement of human-computer interactions. *SoutheastCon , 2005. Proceedings. IEEE*, pages 415–420, 2005.

[87] Philippe Zimmermann, Sissel Guttormsen, Brigitta Danuser, and Patrick Gomez. Affective computing—a rationale for measuring mood with mouse and keyboard. *International journal of occupational safety and ergonomics*, 9(4):539–551, 2003.